#29 attachment 08/945459

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C07K 3/00, 13/00, 15/28, 17/00, C07H 15/12, 17/00, C12N 15/70, 15/79, 15/00, 7/00, 1/20, 5/16, 1/21

A1

(11) International Publication Number: WO 94/15949

(43) International Publication Date: 21 July 1994 (21.07.94)

(21) International Application Number: PCT/US94/00657

(22) International Filing Date: 12 January 1994 (12.01.94)

(30) Priority Data:
08/003,144 12 January 1993 (12.01.93) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/003,144 (CIP)
Filed on 12 January 1993 (12.01.93)

(71) Applicant *(for all designated States except US)*: JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE [US/US]; 720 Rutland Avenue, Baltimore, MD 21205 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LEE, Se-Jin [US/US]; 6711 Chokeberry Road, Baltimore, MD 21209 (US). HUYNH, Thanh [US/US]; 5510 South Bend Road, Baltimore, MD 21209 (US).

(74) Agent: WETHERELL, John, R., Jr.; Spensley Horn Jubas & Lubitz, 5th floor, 1880 Century Park East, Los Angeles, CA 90067 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: GROWTH DIFFERENTIATION FACTOR-5

(57) Abstract

Growth differentiation factor-5 (GDF-5) is disclosed along with its polynucleotide sequence and amino acid sequence. Also disclosed are diagnostic and therapeutic methods of using the GDF-5 polypeptide and polynucleotide sequences.

GROWTH DIFFERENTIATION FACTOR-5

This application is a continuation-in-part application of U.S. Serial No. 08/003,144, filed January 12, 1993.

BACKGROUND OF THE INVENTION

1. *Field of the Invention*

The invention relates generally to growth factors and specifically to a new member of the transforming growth factor beta (TGF-β) superfamily, which is denoted, growth differentiation factor-5 (GDF-5).

2. *Description of Related Art*

The transforming growth factor β (TGF-β) superfamily encompasses a group of structurally-related proteins which affect a wide range of differentiation processes during embryonic development. The family includes, Mullerian inhibiting substance (MIS), which is required for normal male sex development (Behringer *et al.*, *Nature* 345:167, 1990), *Drosophila* decapentaplegic (DPP) gene product, which is required for dorsal-ventral axis formation and morphogenesis of the imaginal disks (Padgett, *et al.*, *Nature*, 325:81-84, 1987), the *Xenopus* Vg-1 gene product, which localizes to the vegetal pole of eggs ((Weeks, *et al.*, *Cell*, 51:861-867, 1987), the activins (Mason, *et al.*, *Biochem, Biophys. Res. Commun.*, 135:957-964, 1986), which can induce the formation of mesoderm and anterior structures in *Xenopus* embryos (Thomsen *et al.*, *Cell* 63:485, 1990), and the bone morphogenetic proteins (BMPs, osteogenin, OP-1) which can induce *de novo* cartilage and bone formation (Sampath, *et al.*, *J. Biol. Chem.* 265:13198, 1990). The TGF-βs can influence a variety of differentiation processes, including adipogenesis, myogenesis, chondrogenesis,

hematopoiesis, and epithelial cell differentiation (for review, see Massague, *Cell* 49:437, 1987).

The proteins of the TGF-β family are initially synthesized as a large precursor protein which subsequently undergoes proteolytic cleavage at a cluster of basic residues approximately 110-140 amino acids from the C-terminus. The C-terminal regions of the proteins are all structurally related and the different family members can be classified into distinct subgroups based on the extent of their homology. Although the homologies within particular subgroups range from 70% to 90% amino acid sequence identity, the homologies between subgroups are significantly lower, generally ranging from only 20% to 50%. In each case, the active species appears to be a disulfide-linked dimer of C-terminal fragments. For most of the family members that have been studied, the homodimeric species has been found to be biologically active, but for other family members, like the inhibins (Ling, *et al., Nature* 321:779, 1986) and the TGF-βs (Cheifetz, *et al., Cell,* 48:409, 1987), heterodimers have also been detected, and these appear to have different biological properties than the respective homodimers.

Identification of new factors that are tissue-specific in their expression pattern will provide a greater understanding of that tissue's development and function.

SUMMARY OF THE INVENTION

The present invention provides a cell growth and differentiation factor, GDF-5, a polynucleotide sequence which encodes the factor and antibodies which are immunoreactive with the factor. This factor appears to relate to various cell proliferative disorders, especially those involving the uterus, such as endometriosis and uterine tumors, and those involving skeletal tissues.

Thus, in one embodiment, the invention provides a method for detecting a cell proliferative disorder of uterine origin and which is associated with GDF-5. In another embodiment, the invention provides a method of treating a cell proliferative disorder associated with expression of GDF-5, by suppressing or enhancing GDF-5 activity.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGURE 1A shows expression of GDF-5 mRNA in adult tissues.

FIGURE 1B shows expression of GDF-5 mRNA in embryonic tissues.

FIGURE 2 shows nucleotide and predicted amino acid sequence of GDF-5. The putative tetrabasic processing sites are denoted by stippled boxes.

FIGURE 3A shows the alignment of the C-terminal sequences of GDF-5 with other members of the TGF-β family. The conserved cysteine residues are shaded. Dashes denote gaps introduced in order to maximize alignment.

FIGURE 3B shows alignment of GDF-5, GDF-6 and GDF-7 C-terminal amino acids.

FIGURE 4 shows amino acid homologies among the different members of the TGF-β superfamily. Numbers represent percent amino acid identities between each pair calculated from the first conserved cysteine to the C-terminus. Boxes represent homologies among highly-related members within particular subgroups.

FIGURE 5 shows shows the expression of GDF-5 in limb mesenchyme of day 12.5 p.c. mouse embryos. Bright field (FIGURE 5a, 5d) and dark field (FIGURE 5b, 5c, 5e, 5f) photomicrographs of transverse (FIGURE 5a-c) and sagittal (FIGURE 5d-f) sections, showing views through forelimb and posterior end of embryo, respectively, after hybridization with ^{35}S-labelled GDF-5 antisense strand (FIGURE 5a,b,d,e) or sense strand control (FIGURE 5c, 5f) probes. Anterior (A), posterior (P), dorsal (D) and ventral (V) orientations are indicated.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a growth and differentiation factor, GDF-5 and a polynucleotide sequence encoding GDF-5. Unlike other members of the TGF-β superfamily, GDF-5 expression is highly tissue specific, being expressed in cells primarily in uterine tissue and skeletal tissue. In one embodiment, the invention provides a method for detection of a cell proliferative disorder of the uterus or skeletal tissue such as bone or cartilage, which is associated with GDF-5 expression. In another embodiment, the invention provides a method for treating a cell proliferative disorder associated with expression of GDF-5 by using an agent which suppresses or enhances GDF-5 activity.

The TGF-β superfamily consists of multifunctionally polypeptides that control proliferation, differentiation, and other functions in many cell types. Many of the peptides have regulatory effects, both positive and negative, on other peptide growth factors. The structural homology between the GDF-5 protein of this invention and the members of the TGF-β family, indicates that GDF-5 is a new member of the family of growth and differentiation factors. Based on the known activities of many of the other members, it can be expected that GDF-5 will also possess biological activities that will make it useful as a diagnostic and therapeutic reagent.

The expression of GDF-5 in the uterus suggests a variety of applications using the polypeptide, polynucleotide, and antibodies of the invention, related to contraception, fertility, pregnancy, and cell proliferative diseases. Abnormally low levels of the factor may be indicative of impair d function in the uterus while abnormally high levels may be indicative of hypertrophy, hyperplasia, or the presence of ectopic tissue. Hence, GDF-5 may be useful in detecting not only primary and metastatic neoplasms of uterine origin but in detecting

diseases such as endometriosis as well. In addition, GDF-5 may also be useful as an indicator of developmental anomalies in prenatal screening procedures.

The expression of GDF-5 during embryogenesis and specifically in the precartilaginous mesenchyme associated with early bone formation in the limbs, suggests a variety of applications using the polypeptide, polynucleotide, and antibodies of the invention, related to skeletal development, cartilage differentiation, and cell proliferative diseases. Abnormally low or high levels of GDF-5 may be indicative of various bone dysplasias such as epiphyseal, physeal (growth plate), metaphyseal and diaphyseal hypo- and hyperplasias. Examples of such diseases which may be diagnosed and/or treated rising GDF-5 polynucleotides and antibodies include: spondyloepithyseal dysplasia, dysplasia epiphysialis hemimelica, achondroplasia, metaphyseal dysostosis, hyperchondroplasia, enchondromatosis, hypophosphatasia, osteopetrosis, craniometaphyseal dysplasia, osteogenesis imperfecta, idiopathic osteoporosis, Engelman's disease and hyperphosphatasia (See Harrison's Principles of Internal Medicine, McGraw-Hill Book Co., N.Y., 1987, Chpt. 339).

Several members of the TGF-β superfamily possess activities suggesting possible applications for the treatment of cell proliferative disorders, such as cancer. In particular, TGF-β has been shown to be potent growth inhibitor for a variety of cell types (Massague, *Cell* 49:437, 1987), MIS has been shown to inhibit the growth of human endometrial carcinoma tumors in nude mice (Donahoe, *et al., Ann. Surg.* 194:472, 1981), and inhibin α has been shown to suppress the development of tumors both in the ovary and in the testis (Matzuk, *et al., Nature*, 360:313, 1992). GDF-5 may have a similar activity and may therefore be useful as an anti-proliferative agent, such as for the treatment of endometrial cancer or endometriosis.

Many of the members of the TGF-β family are also important mediators of tissue repair. TGF-β has been shown to have marked effects on the formation of collagen and causes of striking angiogenic response in the newborn mouse (Roberts, *et al., Proc. Natl. Acad. Sci., USA* 83:4167, 1986). The BMP's can induce new bone growth and are effective for the treatment of fractures and other skeletal defects (Glowacki, *et al., Lancet*, 1:959, 1981; Ferguson, *et al., Clin. Orthoped. Relat. Res.*, 227:265, 1988; Johnson, *et al., Clin Orthoped. Relat. Res.*, 230:257, 1988). Sequence homology and expression data together suggest that GDF-5 may have similar activities and may be useful in repair of tissue injury caused by trauma or burns for example.

GDF-5 may play a role in regulation of the menstrual cycle or regulation of uterine function during pregnancy, and therefore, GDF-5, anti-GDF-5 antibodies, or antisense polynucleotides may be useful either in contraceptive regimens, in enhancing the success of *in vitro* fertilization procedures, or in preventing premature labor.

The term "substantially pure" as used herein refers to GDF-5 which is substantially free of other proteins, lipids, carbohydrates or other materials with which it is naturally associated. One skilled in the art can purify GDF-5 using standard techniques for protein purification. The substantially pure polypeptide will yield a single major band on a non-reducing polyacrylamide gel. The purity of the GDF-5 polypeptide can also be determined by amino-terminal amino acid sequence analysis. GDF-5 polypeptide includes functional fragments of the polypeptide, as long as the activity of GDF-5 remains. Smaller peptides containing the biological activity of GDF-5 ar included in the invention.

The invention provides polynucleotides encoding the GDF-5 protein. These polynucleotides include DNA, cDNA and RNA sequences which encode GDF-5. It is understood that all polynucleotides encoding all or a portion of GDF-5 are

also included herein, as long as they encode a polypeptide with GDF-5 activity. Such polynucleotides include naturally occurring, synthetic, and intentionally manipulated polynucleotides. For example, GDF-5 polynucleotide may be subjected to site-directed mutagenesis. The polynucleotide sequence for GDF-5 also includes antisense sequences. The polynucleotides of the invention include sequences that are degenerate as a result of the genetic code. There are 20 natural amino acids, most of which are specified by more than one codon. Therefore, all degenerate nucleotide sequences are included in the invention as long as the amino acid sequence of GDF-5 polypeptide encoded by the nucleotide sequence is functionally unchanged.

Specifically disclosed herein is a cDNA sequence for GDF-5 which is 2329 base pairs in length and contains an open reading frame beginning with a methionine codon at nucleotide 322. The encoded polypeptide is 495 amino acids in length with a molecular weight of about 54.9 K, as determined by nucleotide sequence analysis. The GDF-5 sequence contains a core of hydrophobic amino acids near the N-terminus, suggestive of a signal sequence for secretion. GDF-5 contains one potential N-glycosylation sites at amino acid 183 and two putative tetrabasic proteolytic processing sites RRKRR and KR-at amino acids 371-375 and amino acids 384-385. Cleavage of the precursor at these sites would generate mature C-terminal fragments of 120 or 110 amino acids in length with predicted molecular weights of 13.6K and 12.5K, respectively.

GDF-5 contains all of the highly conserved residues present in other family members, including the seven cysteine residues with their characteristic spacing. Among the known family members, GDF-5 is most highly related to BMP-2 and BMP-4 in the C-terminal portion of the molecule (57% amino acid sequence identity calculated from the first conserved cysteine).

Minor modifications of the recombinant GDF-5 primary amino acid sequence may result in proteins which have substantially equivalent activity as compared to the GDF-5 polypeptide described herein. Such modifications may be deliberate, as by site-directed mutagenesis, or may be spontaneous. All of the polypeptides produced by these modifications are included herein as long as the biological activity of GDF-5 still exists. Further, deletion of one or more amino acids can also result in a modification of the structure of the resultant molecule without significantly altering its biological activity. This can lead to the development of a smaller active molecule which would have broader utility. For example, one can remove amino or carboxy terminal amino acids which are not required for GDF-5 biological activity.

The nucleotide sequence encoding the GDF-5 polypeptide of the invention includes the disclosed sequence and conservative variations thereof. The term "conservative variation" as used herein denotes the replacement of an amino acid residue by another, biologically similar residue. Examples of conservative variations include the substitution of one hydrophobic residue such as isoleucine, valine, leucine or methionine for another, or the substitution of one polar residue for another, such as the substitution of arginine for lysine, glutamic for aspartic acids, or glutamine for asparagine, and the like. The term "conservative variation" also includes the use of a substituted amino acid in place of an unsubstituted parent amino acid provided that antibodies raised to the substituted polypeptide also immunoreact with the unsubstituted polypeptide.

DNA sequences of the invention can be obtained by several methods. For example, the DNA can be isolated using hybridization techniques which are well known in the art. These include, but are not limited to: 1) hybridization of genomic or cDNA libraries with probes to detect homologous nucleotide

sequences and 2) antibody screening of expression libraries to detect cloned DNA fragments with shared structural features.

Preferably the GDF-5 polynucleotide of the invention is derived from a mammalian organism, and most preferably from a mouse, rat, or human. Screening procedures which rely on nucleic acid hybridization make it possible to isolate any gene sequence from any organism, provided the appropriate probe is available. Oligonucleotide probes, which correspond to a part of the sequence encoding the protein in question, can be synthesized chemically. This requires that short, oligopeptide stretches of amino acid sequence must be known. The DNA sequence encoding the protein can be deduced from the genetic code, however, the degeneracy of the code must be taken into account. It is possible to perform a mixed addition reaction when the sequence is degenerate. This includes a heterogeneous mixture of denatured double-stranded DNA. For such screening, hybridization is preferably performed on either single-stranded DNA or denatured double-stranded DNA. Hybridization is particularly useful in the detection of cDNA clones derived from sources where an extremely low amount of mRNA sequences relating to the polypeptide of interest are present. In other words, by using stringent hybridization conditions directed to avoid non-specific binding, it is possible, for example, to allow the autoradiographic visualization of a specific cDNA clone by the hybridization of the target DNA to that single probe in the mixture which is its complete complement (Wallace, *et al., Nucl. Acid Res.*, 9:879, 1981).

The development of specific DNA sequences encoding GDF-5 can also be obtained by: 1) isolation of double-stranded DNA sequences from the genomic DNA; 2) chemical manufacture of a DNA sequence to provide the necessary codons for the polypeptide of interest; and 3) *in vitro* synthesis of a double-stranded DNA sequence by reverse transcription of mRNA isolated from a

eukaryotic donor cell. In the latter case, a double-stranded DNA complement of mRNA is eventually formed which is generally referred to as cDNA.

Of the three above-noted methods for developing specific DNA sequences for use in recombinant procedures, the isolation of genomic DNA isolates is the least common. This is especially true when it is desirable to obtain the microbial expression of mammalian polypeptides due to the presence of introns.

The synthesis of DNA sequences is frequently the method of choice when the entire sequence of amino acid residues of the desired polypeptide product is known. When the entire sequence of amino acid residues of the desired polypeptide is not known, the direct synthesis of DNA sequences is not possible and the method of choice is the synthesis of cDNA sequences. Among the standard procedures for isolating cDNA sequences of interest is the formation of plasmid- or phage-carrying cDNA libraries which are derived from reverse transcription of mRNA which is abundant in donor cells that have a high level of genetic expression. When used in combination with polymerase chain reaction technology, even rare expression products can be cloned. In those cases where significant portions of the amino acid sequence of the polypeptide are known, the production of labeled single or double-stranded DNA or RNA probe sequences duplicating a sequence putatively present in the target cDNA may be employed in DNA/DNA hybridization procedures which are carried out on cloned copies of the cDNA which have been denatured into a single-stranded form (Jay *et al., Nucl. Acid Res.* 11:2325, 1983).

A cDNA expression library, such as lambda gt11, can be screened indirectly for GDF-5 peptides having at least one epitope, using antibodies specific for GDF-5. Such antibodies can be either polyclonally or monoclonally derived

and used to detect expression product indicative of the presence of GDF-5 cDNA.

DNA sequences encoding GDF-5 can be expressed *in vitro* by DNA transfer into a suitable host cell. "Host cells" are cells in which a vector can be propagated and its DNA expressed. The term also includes any progeny of the subject host cell. It is understood that all progeny may not be identical to the parental cell since there may be mutations that occur during replication. However, such progeny are included when the term "host cell" is used. Methods of stable transfer, meaning that the foreign DNA is continuously maintained in the host, are known in the art.

In the present invention, the GDF-5 polynucleotide sequences may be inserted into a recombinant expression vector. The term "recombinant expression vector" refers to a plasmid, virus or other vehicle known in the art that has been manipulated by insertion or incorporation of the GDF-5 genetic sequences. Such expression vectors contain a promoter sequence which facilitates the efficient transcription of the inserted genetic sequence of the host. The expression vector typically contains an origin of replication, a promoter, as well as specific genes which allow phenotypic selection of the transformed cells. Vectors suitable for use in the present invention include, but are not limited to the T7-based expression vector for expression in bacteria (Rosenberg et al., *Gene* 56:125, 1987), the pMSXND expression vector for expression in mammalian cells (Lee and Nathans, *J. Biol. Chem.* 263:3521, 1988) and baculovirus-derived vectors for expression in insect cells. The DNA segment can be present in the vector operably linked to regulatory elements, for example, a promoter (e.g., T7, metallothionein I, or polyhedrin promoters).

Polynucleotide sequences encoding GDF-5 can be expressed in either prokaryotes or eukaryotes. Hosts can include microbial, yeast, insect and mammalian organisms. Methods of expressing DNA sequences having eukaryotic or viral sequences in prokaryotes are well known in the art. Biologically functional viral and plasmid DNA vectors capable of expression and replication in a host are known in the art. Such vectors are used to incorporate DNA sequences of the invention.

Transformation of a host cell with recombinant DNA may be carried out by conventional techniques as are well known to those skilled in the art. Where the host is prokaryotic, such as *E. coli*, competent cells which are capable of DNA uptake can be prepared from cells harvested after exponential growth phase and subsequently treated by the $CaCl_2$ method using procedures well known in the art. Alternatively, $MgCl_2$ or RbCl can be used. Transformation can also be performed after forming a protoplast of the host cell if desired.

When the host is a eukaryote, such methods of transfection of DNA as calcium phosphate co-precipitates, conventional mechanical procedures such as microinjection, electroporation, insertion of a plasmid encased in liposomes, or virus vectors may be used. Eukaryotic cells can also be cotransformed with DNA sequences encoding the GDF-5 of the invention, and a second foreign DNA molecule encoding a selectable phenotype, such as the herpes simplex thymidine kinase gene. Another method is to use a eukaryotic viral vector, such as simian virus 40 (SV40) or bovine papilloma virus, to transiently infect or transform eukaryotic cells and express the protein. (see for example, *Eukaryotic Viral Vectors*, Cold Spring Harbor Laboratory, Gluzman ed., 1982).

Isolation and purification of microbial expressed polypeptide, or fragments thereof, provided by the invention, may be carried out by conventional means including preparative chromatography and immunological separations involving monoclonal or polyclonal antibodies.

The invention includes antibodies immunoreactive with GDF-5 polypeptide or functional fragments thereof. Antibody which consists essentially of pooled monoclonal antibodies with different epitopic specificities, as well as distinct monoclonal antibody preparations are provided. Monoclonal antibodies are made from antigen containing fragments of the protein by methods well known to those skilled in the art (Kohler, et al., Nature, 256:495, 1975). The term antibody as used in this invention is meant to include intact molecules as well as fragments thereof, such as Fab and $F(ab')_2$, which are capable of binding an epitopic determinant on GDF-5.

The term "cell-proliferative disorder" denotes malignant as well as non-malignant cell populations which often appear to differ from the surrounding tissue both morphologically and genotypically. The GDF-5 polynucleotide that is an antisense molecule is useful in treating cell proliferative disorders of the various organ systems, particularly, for example, the uterus or skeletal system. Cell proliferative disorders of the skeletal system include those disorders of bone cells and cartilage as described above. Essentially, any disorder involving cells that are normally responsive to GDF-5 could be considered susceptible to treatment with a GDF-5 suppressing reagent.

The invention provides a method for detecting a cell proliferative disorder of the uterus or skeletal system (e.g., bone, cartilage) which comprises contacting an anti-GDF-5 antibody with a cell suspected of having a GDF-5 associated disorder and detecting binding to the antibody. The antibody reactive with GDF-5 is labeled with a compound which allows detection of binding to GDF-5.

For purposes of the invention, an antibody specific for GDF-5 polypeptide may be used to detect the level of GDF-5 in biological fluids and tissues. Any specimen containing a detectable amount of antigen can be used. A preferred sample of this invention is tissue of uterine origin, specifically endometrial tissue or skeletal tissue such as bone and cartilage. The level of GDF-5 in the suspect cell can be compared with the level in a normal cell to determine whether the subject has a GDF-5-associated cell proliferative disorder. Preferably the subject is human.

The antibodies of the invention can be used in any subject in which it is desirable to administer *in vitro* or *in vivo* immunodiagnosis or immunotherapy. The antibodies of the invention are suited for use, for example, in immunoassays in which they can be utilized in liquid phase or bound to a solid phase carrier. In addition, the antibodies in these immunoassays can be detectably labeled in various ways. Examples of types of immunoassays which can utilize antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of the antigens using the antibodies of the invention can be done utilizing immunoassays which are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

The antibodies of the invention can be bound to many different carriers and used to detect the presence of an antigen comprising the polypeptide of the invention. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled

in the art will know of other suitable carriers for binding antibodies, or will be able to ascertain such, using routine experimentation.

There are many different labels and methods of labeling known to those of ordinary skill in the art. Examples of the types of labels which can be used in the present invention include enzymes, radioisotopes, fluorescent compounds, colloidal metals, chemiluminescent compounds, phosphorescent compounds, and bioluminescent compounds. Those of ordinary skill in the art will know of other suitable labels for binding to the antibody, or will be able to ascertain such, using routine experimentation.

Another technique which may also result in greater sensitivity consists of coupling the antibodies to low molecular weight haptens. These haptens can then be specifically detected by means of a second reaction. For example, it is common to use such haptens as biotin, which reacts with avidin, or dinitrophenyl, puridoxal, and fluorescein, which can react with specific anti-hapten antibodies.

In using the monoclonal antibodies of the invention for the *in vivo* detection of antigen, the detectably labeled antibody is given a dose which is diagnostically effective. The term "diagnostically effective" means that the amount of detectably labeled monoclonal antibody is administered in sufficient quantity to enable detection of the site having the antigen comprising a polypeptide of the invention for which the monoclonal antibodies are specific.

The concentration of detectably labeled monoclonal antibody which is adminstered should be sufficient such that the binding to those cells having the polypeptide is detectable compared to the background. Further, it is desirable that the detectably labeled monoclonal antibody be rapidly cleared from the circulatory system in order to give the best target-to-background signal ratio.

As a rule, the dosage of detectably labeled monoclonal antibody for in vivo diagnosis will vary depending on such factors as age, sex, and extent of disease of the individual. Such dosages may vary, for example, depending on whether multiple injections are given, antigenic burden, and other factors known to those of skill in the art.

For *in vivo* diagnostic imaging, the type of detection instrument available is a major factor in selecting a given radioisotope. The radioisotope chosen must have a type of decay which is detectable for a given type of instrument. Still another important factor in selecting a radioisotope for *in vivo* diagnosis is that deleterious radiation with respect to the host is minimized. Ideally, a radioisotope used for *in vivo* imaging will lack a particle emission, but produce a large number of photons in the 140-250 keV range, which may readily be detected by conventional gamma cameras.

For *in vivo* diagnosis radioisotopes may be bound to immunoglobulin either directly or indirectly by using an intermediate functional group. Intermediate functional groups which often are used to bind radioisotopes which exist as metallic ions to immunoglobulins are the bifunctional chelating agents such as diethylenetriaminepentacetic acid (DTPA) and ethylenediaminetetraacetic acid (EDTA) and similar molecules. Typical examples of metallic ions which can be bound to the monoclonal antibodies of the invention are ^{111}In, ^{97}Ru, ^{67}Ga, ^{68}Ga, ^{72}As, ^{89}Zr, and ^{201}Tl.

The monoclonal antibodies of the invention can also be labeled with a paramagnetic isotope for purposes of *in vivo* diagnosis, as in magnetic resonance imaging (MRI) or electron spin resonance (ESR). In general, any conventional method for visualizing diagnostic imaging can be utilized. Usually gamma and positron emitting radioisotopes are used for camera imaging and

paramagnetic isotopes for MRI. Elements which are particularly useful in such techniques include ^{157}Gd, ^{55}Mn, ^{162}Dy, ^{52}Cr, and ^{56}Fe.

The monoclonal antibodies of the invention can be used *in vitro* and *in vivo* to monitor the course of amelioration of a GDF-5-associated disease in a subject. Thus, for example, by measuring the increase or decrease in the number of cells expressing antigen comprising a polypeptide of the invention or changes in the concentration of such antigen present in various body fluids and tissues, it would be possible to determine whether a particular therapeutic regimen aimed at ameliorating the GDF-5-associated disease is effective. The term "ameliorate" denotes a lessening of the detrimental effect of the GDF-5-associated disease in the subject receiving therapy.

The present invention identifies a nucleotide sequence that can be expressed in an altered manner as compared to expression in a normal cell, therefore it is possible to design appropriate therapeutic or diagnostic techniques directed to this sequence. Thus, where a cell-proliferative disorder is associated with the expression of GDF-5, nucleic acid sequences that interfere with GDF-5 expression at the translational level can be used. This approach utilizes, for example, antisense nucleic acid and ribozymes to block translation of a specific GDF-5 mRNA, either by masking that mRNA with an antisense nucleic acid or by cleaving it with a ribozyme.

Antisense nucleic acids are DNA or RNA molecules that are complementary to at least a portion of a specific mRNA molecule (Weintraub, *Scientific American*, 262:40, 1990). In the cell, the antisense nucleic acids hybridize to the corresponding mRNA, forming a double-stranded molecule. The antisense nucleic acids interfere with the translation of the mRNA, since the cell will not translate a mRNA that is double-stranded. Antisense oligomers of about 15 nucleotides are preferred, since they are easily synthesized and are less likely

to cause problems than larger molecules when introduced into the target GDF-5-producing cell. The use of antisense methods to inhibit the *in vitro* translation of genes is well known in the art (Marcus-Sakura, *Anal.Biochem.*, 172:289, 1988).

Ribozymes are RNA molecules possessing the ability to specifically cleave other single-stranded RNA in a manner analogous to DNA restriction endonucleases. Through the modification of nucleotide sequences which encode these RNAs, it is possible to engineer molecules that recognize specific nucleotide sequences in an RNA molecule and cleave it (Cech, *J.Amer.Med. Assn.*, 260:3030, 1988). A major advantage of this approach is that, because they are sequence-specific, only mRNAs with particular sequences are inactivated.

There are two basic types of ribozymes namely, *tetrahymena*-type (Hasselhoff, *Nature*, 334:585, 1988) and "hammerhead"-type. *Tetrahymena*-type ribozymes recognize sequences which are four bases in length, while "hammerhead"-type ribozymes recognize base sequences 11-18 bases in length. The longer the recognition sequence, the greater the likelihood that the sequence will occur exclusively in the target mRNA species. Consequently, hammerhead-type ribozymes are preferable to *tetrahymena*-type ribozymes for inactivating a specific mRNA species and 18-based recognition sequences are preferable to shorter recognition sequences.

The present invention also provides gene therapy for the treatment of cell proliferative disorders which are mediated by GDF-5 protein. Such therapy would achieve its therapeutic effect by introduction of the GDF-5 antisense polynucleotide into cells having the proliferative disorder. Delivery of antisense GDF-5 polynucleotide can be achieved using a recombinant expression vector such as a chimeric virus or a colloidal dispersion system. Especially preferred

for therapeutic delivery of antisense sequences is the use of targeted liposomes.

Various viral vectors which can be utilized for gene therapy as taught herein include adenovirus, herpes virus, vaccinia, or, preferably, an RNA virus such as a retrovirus. Preferably, the retroviral vector is a derivative of a murine or avian retrovirus. Examples of retroviral vectors in which a single foreign gene can be inserted include, but are not limited to: Moloney murine leukemia virus (MoMuLV), Harvey murine sarcoma virus (HaMuSV), murine mammary tumor virus (MuMTV), and Rous Sarcoma Virus (RSV). A number of additional retroviral vectors can incorporate multiple genes. All of these vectors can transfer or incorporate a gene for a selectable marker so that transduced cells can be identified and generated. By inserting a GDF-5 sequence of interest into the viral vector, along with another gene which encodes the ligand for a receptor on a specific target cell, for example, the vector is now target specific. Retroviral vectors can be made target specific by inserting, for example, a polynucleotide encoding a sugar, a glycolipid, or a protein. Preferred targeting is accomplished by using an antibody to target the retroviral vector. Those of skill in the art will know of, or can readily ascertain without undue experimentation, specific polynucleotide sequences which can be inserted into the retroviral genome to allow target specific delivery of the retroviral vector containing the GDF-5 antisense polynucleotide.

Since recombinant retroviruses are defective, they require assistance in order to produce infectious vector particles. This assistance can be provided, for example, by using helper cell lines that contain plasmids encoding all of the structural genes of the retrovirus under the control of regulatory sequences within the LTR. These plasmids are missing a nucleotide sequence which enables the packaging mechanism to recognize an RNA transcript for encapsidation. Helper cell lines which have deletions of the packaging signal

include, but are not limited to ψ2, PA317 and PA12, for example. These cell lines produce empty virions, since no genome is packaged. If a retroviral vector is introduced into such cells in which the packaging signal is intact, but the structural genes are replaced by other genes of interest, the vector can be packaged and vector virion produced.

Alternatively, NIH 3T3 or other tissue culture cells can be directly transfected with plasmids encoding the retroviral structural genes *gag, pol* and *env*, by conventional calcium phosphate transfection. These cells are then transfected with the vector plasmid containing the genes of interest. The resulting cells release the retroviral vector into the culture medium.

Another targeted delivery system for GDF-5 antisense polynucleotides is a colloidal dispersion system. Colloidal dispersion systems include macromolecule complexes, nanocapsules, microspheres, beads, and lipid-based systems including oil-in-water emulsions, micelles, mixed micelles, and liposomes. The preferred colloidal system of this invention is a liposome. Liposomes are artificial membrane vesicles which are useful as delivery vehicles *in vitro* and *in vivo*. It has been shown that large unilamellar vesicles (LUV), which range in size from 0.2-4.0 μm can encapsulate a substantial percentage of an aqueous buffer containing large macromolecules. RNA, DNA and intact virions can be encapsulated within the aqueous interior and be delivered to cells in a biologically active form (Fraley, *et al., Trends Biochem. Sci.*, 6:77, 1981). In addition to mammalian cells, liposomes have been used for delivery of polynucleotides in plant, yeast and bacterial cells. In order for a liposome to be an efficient gene transfer vehicle, the following characteristics should be present: (1) encapsulation of the genes of interest at high efficiency while not compromising their biological activity; (2) preferential and substantial binding to a target cell in comparison to non-target cells; (3) delivery of the aqueous contents of the vesicle to the target cell cytoplasm at high efficiency; and (4)

accurate and effective expression of genetic information (Mannino, *et al.*, *Biotechniques*, 6:682, 1988).

The composition of the liposome is usually a combination of phospholipids, particularly high-phase-transition-temperature phospholipids, usually in combination with steroids, especially cholesterol. Other phospholipids or other lipids may also be used. The physical characteristics of liposomes depend on pH, ionic strength, and the presence of divalent cations.

Examples of lipids useful in liposome production include phosphatidyl compounds, such as phosphatidylglycerol, phosphatidylcholine, phosphatidylserine, phosphatidylethanolamine, sphingolipids, cerebrosides, and gangliosides. Particularly useful are diacylphosphatidylglycerols, where the lipid moiety contains from 14-18 carbon atoms, particularly from 16-18 carbon atoms, and is saturated. Illustrative phospholipids include egg phosphatidylcholine, dipalmitoylphosphatidylcholine and distearoylphosphatidylcholine.

The targeting of liposomes can be classified based on anatomical and mechanistic factors. Anatomical classification is based on the level of selectivity, for example, organ-specific, cell-specific, and organelle-specific. Mechanistic targeting can be distinguished based upon whether it is passive or active. Passive targeting utilizes the natural tendency of liposomes to distribute to cells of the reticulo-endothelial system (RES) in organs which contain sinusoidal capillaries. Active targeting, on the other hand, involves alteration of the liposome by coupling the liposome to a specific ligand such as a monoclonal antibody, sugar, glycolipid, or protein, or by changing the composition or size of the liposome in order to achieve targeting to organs and cell types other than the naturally occurring sites of localization.

The surface of the targeted delivery system may be modified in a variety of ways. In the case of a liposomal targeted delivery system, lipid groups can be incorporated into the lipid bilayer of the liposome in order to maintain the targeting ligand in stable association with the liposomal bilayer. Various linking groups can be used for joining the lipid chains to the targeting ligand.

The following examples are intended to illustrate but not limit the invention. While they are typical of those that might be used, other procedures known to those skilled in the art may alternatively be used.

EXAMPLE 1

IDENTIFICATION AND ISOLATION OF A NOVEL TGF-ß FAMILY MEMBER

To identify a new member of the TGF-β superfamily, degenerate oligonucleotides were designed which corresponded to two conserved regions among the known family members: one region spanning the two tryptophan residues conserved in all family members except MIS and the other region spanning the invariant cysteine residues near the C-terminus. These primers were used for polymerase chain reactions on mouse genomic DNA followed by subcloning the PCR products using restriction sites placed at the 5' ends of the primers, picking individual *E. coli* colonies carrying these subcloned inserts, and using a combination of random sequencing and hybridization analysis to eliminate known members of the superfamily.

GDF-5 was identified by polymerase chain reaction (PCR) using mouse genomic DNA with the following primers:

SJL136:5'-CCGGAATTCGG(G/A/T/C)TGGGA(G/A)(A/C)G(G/A/T/C)TGG(G/A)T (G/A/T/C)(G/A)T-3' (SEQUENCE ID NO. 1)

SJL121:5'-CCGGAATTC(G/A)CAICC(G/A)CA(T/C)TC(G/A)TCIACIACCAT(G/A)TC(T/C)TC(G/A)TA-3' (SEQUENCE ID NO. 2)

SJL 136 corresponds to the amino acid sequence GWE(R/S)W(V/I/M)(V/I/M), (SEQUENCE ID NO. 3) and the complement of SJL 121 corresponds to the amino acid sequence YEDMVVDECGC (SEQUENCE ID NO. 4). Both oligonucleotide sets were designed to contain an EcoRI restriction site at the 5'end to facillitate subcloning. PCR was carried out for 40 cycles at 94°C for 1', 50°C for 2' and 72°C for 3.5'.

Human GDF-5 was isolated by PCR using human genomic DNA with the following primers:

SJL 141: 5'-CCGGAATTCGGITGG(G/C/A)A(G/A/T/C)(A/G)A(T/C)TGG(A/G)TI(A/G)TI(T/G)CICC-3' (SEQUENCE ID NO. 5)

SJL145:5'-CCGGAATTC(G/A)CAI(G/C)C(G/A)CAIG(C/A)(G/A/T/C)TCIACI(G/A)(T/C)CAT-3' (SEQUENCE ID NO. 6)

SSJL 141 corresponds to the amino acid sequence GW(H/Q/N/K/D/E)(D/N)W-(V/I/M)(V/I/M)(A/S)P (SEQUENCE ID NO. 7) and the complement of SJL 145 corresponds to the amino acid sequence M(V/I/M/T/A)V(D/E)(A/S)C(G/A)C (SEQUENCE ID NO. 8). Both the oligonucleotide sets were designed to contain an EcoRI restriction site at the 5' end to facilitate subcloning. PCR was carried out for 40 cycles at 94°C for 1 min., 50°C for 2 min., and 72°C for 2 min. Partial sequence analysis of the human PCR product revealed no predicted amino acid differences between mouse and human GDF-5.

PCR products of approximately 280 bp were gel-purified, digested with Eco RI, gel-purified again, and subcloned in the Bluescript vector (Stratagene, San Diego, CA). Bacterial colonies carrying individual subclones were picked into 96 well microtiter plates, and multiple replicas were prepared by plating the cells onto nitrocellulose. The replicate filters were hybridized to probes

representing known members of the family, and DNA was prepared from non-hybridizing colonies for sequence analysis.

RNA isolation and Northern analysis were carried out as described previously (Lee,S.J., *Mol. Endocrinol.* 4:1034, 1990). An oligo dT-primed cDNA library was prepared from 2.5-3 μg of 12.5 day gestation CD-1 mouse embryo poly A-selected RNA in the lambda ZAP II vector according to the instructions provided by Stratagene. The library was amplified prior to screening. Filters were hybridized as described previously (Lee, S.-J., *Proc. Natl. Acad. Sci. USA.*, 88:4250-4254, 1991). DNA sequencing of both strands was carried out using the dideoxy chain termination method (Sanger, *et al., Proc. Natl. Acad. Sci., USA* 74:5463-5467, 1977) and a combination of the S1 nuclease-/exonuclease III strategy (Henikoff, S., *Gene*, 28:351-359, 1984) and synthetic oligonucleotide primers.

EXAMPLE 2

EXPRESSION PATTERN AND SEQUENCE OF GDF-5

To determine the expression pattern of GDF-5, RNA samples prepared from a variety of adult tissues were screened by Northern analysis. RNA isolation and Northern analysis were carried out as described previously (Lee, S.J., *Mol. Endocrinol.*, 4:1034, 1990). Five micrograms of twice polyA-selected RNA prepared from each tissue were electrophoresed on formaldehyde gels, blotted and probed with GDF-5. As shown in Figure 1A, the GDF-5 probe detected an approximately 2.5 kb mRNA expressed primarily in the uterus and at lower levels in other adult tissues in the mouse, including placenta, brain, thymus, lung, kidney, and adrenal gland. The GDF-5 probe also detected a larger mRNA in the oviduct. High levels of GDF-5 transcripts were also detected in mouse embryos, particularly at day 12.5 of gestation (FIGURE 1B).

A CD-1 day 12.5 whole mouse embryo cDNA library was constructed in lambda ZAP II and screened with a probe derived from the GDF-5 PCR product. The nucleotide sequence of the longest hybridizing clone is shown in Figure 2. The in-frame termination codons upstream of the putative initiating ATG and the consensus polyadenylation signals are underlined. The poly A tails are not shown. Numbers indicate nucleotide position relative to the 5' end. The 2329 bp sequence contains a long open reading frame beginning with a methionine codon at nucleotide 322 and potentially encoding a protein 495 amino acids in length with a molecular weight of 54.9 K. Like other TGF-β family members, the GDF-5 sequence contains a core of hydrophobic amino acids near the N-terminus suggestive of a signal sequence for secretion. GDF-5 contains a single potential N-glycosylation sites at asparagine residue 183 (denoted by the plain box) and two putative tetrabasic proteolytic processing sites at amino acids 371-375 (denoted by the stippled box) and amino acids 384-385. GDF-5 contains all of the highly conserved residues present in other family members (Figures 3 and 4), including the seven cysteine residues with their characteristic spacing. Among the known mammalian family members, GDF-5 is most highly related to BMP-2 and BMP-4 in the C-terminal portion of the molecule (57% amino acid sequence identity calculated from the first conserved cysteine).

Although the C-terminal portion of GDF-5 clearly shows homology with the other family members, the sequence of GDF-5 is significantly diverged from those of the other family members (Figures 3 and 4). Figure 3 shows the alignment of the C-terminal sequences of GDF-5 with the corresponding regions of human GDF-1 (Lee, *Proc. Natl. Acad. Sci. USA* 88:4250-4254, 1991), human Vgr-1 (Celeste, *et al., Proc. Natl. Acad. Sci. USA* 87:9843-9847, 1990), human OP-1 (Ozkaynak, *et al., EMBO J.* 9:2085-2093, 1990), human BMP-5 (Celeste, *et al., Proc. Natl. Acad. Sci. USA*, 87:9843-9847, 1990), human BMP-3 (Wozney, *et al., Science*, 242:1528-1534, 1988), human MIS (*Cate, et al. Cell*,

45:685-698, 1986), human inhibin α, βA, and βB (Mason, *et al., Biochem, Biophys. Res. Commun.*, 135:957-964, 1986), human TGF-β1 (Derynck, *et al., Nature,* 316:701-705, 1985), humanTGF-β2 (deMartin, *et al., EMBO J.*, 6:3673-3677, 1987), human TGF-β3 (ten Dijke, *et al., Proc. Natl. Acad. Sci. USA,* 85:4715-4719, 1988), chicken TGF-β4 (Jakowlew, *et al., Mol. Endocrinol.* 2:1186-1195, 1988), and *Xenopus* TGF-β5 (Kondaiah, *et al., J. Biol. Chem.* 265:1089-1093, 1990). The conserved cysteine residues are boxed. Dashes denote gaps introduced in order to maximize the alignment.

Figure 4 shows the amino acid homologies among the different members of the TGF-β superfamily. Numbers represent percent amino acid identities between each pair calculated from the first conserved cysteine to the C-terminus. Boxes represent homologies among highly-related members within particular subgroups.

The degree of sequence identify with known family members ranges from a minimum of 24% with inhibin alpha to a maximum of 57% with BMP-2 and BMP-4. GDF-5 shows no significant sequence homology to other family members in the pro-region of the molecule.

EXAMPLE 3

The results in Example 2 show that during the development of the mouse embryo, the expression of GDF-5 begins at approximately day 10.5 post coitum (p.c.) and peaks at day 12.5 p.c., as indicated by the presence of a 2.5 kilobase (kb) major transcript (FIGURE 1B). Of the adult mouse tissues examined, uterus contained the highest level of the 2.5 kb transcript, while low levels were detected in placenta (day 10.5 p.c.), oviduct, brain, thymus, heart, lung, kidney and adrenal gland (FIGURE 1A). In oviduct tissue, the GDF-5

probe also detected a larger transcript of approximately 3.6 kb. GDF-5 transcripts were also detected by Northern blot analysis in femur and calvaria of newborn mice.

In order to characterize in more detail, the expression of GDF-5 in embryonic tissues, ^{35}S-labelled probes synthesized from a portion of the cDNA clone encoding the relatively nonconserved prepro-region were hybridized *in situ* to serial sections of day 12.5 p.c. embryos. Day 12.5 p.c. female CD-1 mouse embryos were fixed and embedded in paraffin as described (Jones, C.M., *et al., Development*, 111:531-542, 1991). ^{35}S-labelled antisense or sense strand RNA probes were synthesized by *in vitro* transcription from a template containing nucleotides 308 through 1446 of the GDF-5 cDNA clone (FIGURE 2). Eight micron sections were hybridized with antisense or sense strand probe at 4 x 10^5 counts per minute/μl essentially as described (Jones, C.M., *et al., supra*) except that the proteinase K and acetic anhydride treatments were omitted, washes in 50% formamide, 2 x SSC, 0.1 M DTT were carried out at 65°C, and the final wash in 0.1 x SSC was carried out at 37°C. Slides were developed after a 4-6 week exposure time with Kodak NTB3 emulsion and were stained with hematoxylin and eosin.

FIGURE 5 shows shows the expression of GDF-5 in limb mesenchyme of day 12.5 p.c. mouse embryos. Bright field (FIGURE 5a, 5d) and dark field (FIGURE 5b, 5c, 5e, 5f) photomicrographs of transverse (FIGURE 5a-c) and sagittal (FIGURE 5d-f) sections, showing views through forelimb and posterior end of embryo, respectively, after hybridization with ^{35}S-labelled GDF-5 antisense strand (FIGURE 5a,b,d,e) or sense strand control (FIGURE 5c, 5f) probes. Serial sections revealed hybridization to be localized to proximal (closed arrows) and distal (open arrows) mesenchyme in the forelimb (FIGURE 5a-c) and hindlimb (FIGURE 5d-f). Anterior (A), posterior (P), dorsal (D) and ventral (V) orientations are indicated.

GDF-5 transcripts were detected in both proximal and distal precartilaginous mesenchyme of the forelimbs and hindlimbs (FIGURE 5). No other major sites of hybridization in the embryo were detected. The development of the long bones of the limbs begins with the condensation of mesenchyme, which differentiates into cartilage-forming cells. Osteogenic cells eventually invade the cartilage matrix and produce a bone matrix which becomes ossified (Rosen, V., *et al., Trends Genet.*, 8:97-102, 1992). In the mouse embryo at 12.5 days p.c., cartilage formation is just beginning in the long bones, and no sign of ossification is yet seen (Kaufman, M.H., *The Atlas of Mouse Development*, Academic Press, Inc., 1992). The peak of GDF-5 expression at this stage (FIGURE 1B) and its primary location in the precartilaginous limb mesenchyme suggest that GDF-5 may affect the production, proliferation, and/or differentiation of the mesenchyme cells.

In addition to GDF-5, two other members of the TGF-β superfamily have been suggested to play a role in limb development. In particular, BMP-2 and BMP-4 are known to be expressed in the apical ectodermal ridge (AER) during mid-gestation at day 10.5 p.c. (Lyons, K.M., *et al., Development*, 109:833-844, 1990; Jones, C.M., *et al., Development*, 111:531-542, 1991). BMP-2 has been shown to inhibit the proliferation of mesenchyme cells in cultured limbs of mid-gestational embryos from which the AER had been removed (Niswander, L., *et al., Nature*, 361:68-71, 1993). Because BMP-2 and BMP-4 are also known to be expressed in limb mesenchyme at day 12.5 p.c. and because the active form of growth factors in this family is generally a disulfied-linked dimer, the possibility exists that homodimers or heterodimers of GDF-5, BMP-2 and BMP-4 may have distinct roles in limb development.

So far, the only bone morphogenetic protein for which mutants have been found is BMP-5, encoded by the mouse short ear locus (Kingsley, D.M., *et al., Cell*, 71:399-419, 1992). Mice homozygous for the short ear mutation, which

causes a range of skeletal defects, have alterations in the size and shape of precartilaginous condensations of mesenchyme (Green, E.L., *et al., J. Morphol.*, 70:1-19, 1942). Skeletal defects of the limbs and digits may be caused by mutations in the mouse gene encoding GDF-5. Like BMP-5, GDF-5 controls particular aspects of skeletal morphology during development.

-- SUMMARY OF SEQUENCES

SEQUENCE ID NO 1 is the nucleotide sequence for the GDF-5 primer, SJL136.

SEQUENCE ID NO 2 is the nucleotide sequence for the GDF-5 primer, SJL121.

SEQUENCE ID NO 3 is the amino acid sequence for the GDF-5 primer, SJL136.

SEQUENCE ID NO 4 is the amino acid sequence for the GDF-5 primer, SJL121.

SEQUENCE ID NO 5 is the nucleotide sequence for the GDF-5 primer, SJL141.

SEQUENCE ID NO 6 is the nucleotide sequence for the GDF-5 primer, SJL145.

SEQUENCE ID NO 7 is the amino acid sequence for the GDF-5 primer, SJL141.

SEQUENCE ID NO 8 is the amino acid sequence for the GDF-5 primer, SJL145.

SEQUENCE ID NO 9 is the nucleotide and deduced amino acid sequence for GDF-5.

SEQUENCE ID NO 10 is the deduced amino acid sequence for GDF-5.

SEQUENCE ID NO 11 is the amino acid sequence for GDF-1.

SEQUENCE ID NO 12 is the amino acid sequence for GDF-3.

SEQUENCE ID NO 13 is the amino acid sequence for GDF-5.

SEQUENCE ID NO 14 is the amino acid sequence for GDF-9.

SEQUENCE ID NO 15 is the amino acid sequence for BMP-2.

SEQUENCE ID NO 16 is the amino acid sequence for GDF-4.

SEQUENCE ID NO 17 is the amino acid sequence for Vgr-1.

SEQUENCE ID NO 18 is the amino acid sequence for Op-1.

SEQUENCE ID NO 19 is the amino acid sequence for BMP-5.

SEQUENCE ID NO 20 is the amino acid sequence for BMP-3.

SEQUENCE ID NO 21 is the amino acid sequence for MIS.

SEQUENCE ID NO 22 is the amino acid sequence for inhibin-α.

SEQUENCE ID NO 23 is the amino acid sequence for inhibin-$\beta\alpha$.

SEQUENCE ID NO 24 is the amino acid sequence for inhibin-$\beta\beta$.

SEQUENCE ID NO 25 is the amino acid sequence for TGF-β1.

SEQUENCE ID NO 26 is the amino acid sequence for TGF-β2.

SEQUENCE ID NO 27 is the amino acid sequence for TGF-β3.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: SE-JIN LEE
HUYNH, THANH

(ii) TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-5

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
(B) STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
(C) CITY: LOS ANGELES
(D) STATE: CALIFORNIA
(E) COUNTRY: US
(F) ZIP: 90067

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT
(B) FILING DATE: 1/12/94
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: WETHERELL, JR. PH.D., JOHN R.
(B) REGISTRATION NUMBER: 31,678
(C) REFERENCE/DOCKET NUMBER: FD3256 CIP OF PD2280

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619/455-5100
(B) TELEFAX: 619-455-5110

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: 136

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..28

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCGGAATTCG GNTGGGARMG NTGGRTNR 28

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: 121

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..42
(D) OTHER INFORMATION: /note= "WHERE "B" OCCURS, B = INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CCGGAATTCR CABCCRCAYT CRTCBACBAC CATRTCYTCR TA 42

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(B) CLONE: 136

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..7
(D) OTHER INFORMATION: /note= "R = Arg, Ser; V = Val, Ileu, Met."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Gly Trp Glu Arg Trp Val Val
1 5

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(B) CLONE: 121

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..11

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Tyr Glu Asp Met Val Val Asp Glu Cys Gly Cys
1 5 10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: 141

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..35
(D) OTHER INFORMATION: /note= "WHERE "B" OCCURS, B = INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CCGGAATTCG GBTGGVANRA YTGGRTBRTB KCBCC 35

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: 145

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..33

(D) OTHER INFORMATION: /note= "WHERE "B" OCCURS, B = INOSINE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCGGAATTCR CABSCRCABG MNTCBACBRY CAT 33

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(B) CLONE: 141

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..9
(D) OTHER INFORMATION: /note= "H = His, Gln, Asn, Lys, Glu, Asp; D = Asp, Asn; V = Val, Ile, Met; A = Glu, Ser. "

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gly Trp His Asp Trp Val Val Ala Pro
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(B) CLONE: 145

(ix) FEATURE:
(A) NAME/KEY: Peptide
(B) LOCATION: 1..8
(D) OTHER INFORMATION: /note= "V = Val, Ile, Met, Thr, Ala; D = Asp, Glu; A = Ala, Ser; G = Gly, ..."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Val Asp Ala Cys Gly Cys
1 5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2329 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vii) IMMEDIATE SOURCE:
(B) CLONE: GDF-5

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 322..1807

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTCAAGCCCT CAGTCAGTTG TGCGGGAGAA AGGGGGCGGT CGGCTTTCTC CTTTCAAGAA 60

CGAGTTATTT TCAGCTGCTG ACTGGAGACG GTGCACGTCT GGACACGGGA GCACTTCCAC 120

TATGGGACTG GATACAGACA CACGCCCGGC GGACTTCAAG ACACTCAGAC TGAGGAGAAA 180

GCCCTGCCTG CTGCTGCTGC TGCTGCTGCT GCCACCGCTG CCTCTGAAGA CCCACTCCTT 240

TCATGGTTTT TCCTGCCAAG CCAGAGGCAC CTTCGCTGCT ACGGCCTTTC TCTGTGGTGT 300

```
CATTCAGCGG CTGGCCAGAG G ATG AGA CTC CCC AAA CTC CTC ACT CTT TTG      351
                        Met Arg Leu Pro Lys Leu Leu Thr Leu Leu
                          1               5                  10

CTG TGG CAC CTG GCT TGG CTG GAC CTG GAA CTC ATC TGC ACT GTG CTG      399
Leu Trp His Leu Ala Trp Leu Asp Leu Glu Leu Ile Cys Thr Val Leu
                  15                  20                  25

GGT GCC CCT GAC TTA GGA CAG AGA ACC CCA GGG GCC AAG CCA GGG TTG      447
Gly Ala Pro Asp Leu Gly Gln Arg Thr Pro Gly Ala Lys Pro Gly Leu
              30                  35                  40

ACC AAA GCG GAG GCC AAG GAG AGG CCA CCC CTG GCC AGG AAT GTC TTT      495
Thr Lys Ala Glu Ala Lys Glu Arg Pro Pro Leu Ala Arg Asn Val Phe
          45                  50                  55

AGG CCA GGG GGT CAT ATC TAT GGT GTG GGG GCC ACC AAT GCC AGG GCC      543
Arg Pro Gly Gly His Ile Tyr Gly Val Gly Ala Thr Asn Ala Arg Ala
      60                  65                  70

AAG GGA AGC TCT GGG CAG ACA CAG GCC AAG AAG GAT GAA CCC AGA AAG      591
Lys Gly Ser Ser Gly Gln Thr Gln Ala Lys Lys Asp Glu Pro Arg Lys
  75                  80                  85                  90

ATG CCC CCC AGA TCC GGT GGC TCT GAA ACC AAG CCA GGA CCC TCT TCC      639
Met Pro Pro Arg Ser Gly Gly Ser Glu Thr Lys Pro Gly Pro Ser Ser
                  95                 100                 105

CAG ACT AGA CAG GCT GCA GCC CGG ACT GTA ACC CCA AAA GGA CAG CTT      687
Gln Thr Arg Gln Ala Ala Ala Arg Thr Val Thr Pro Lys Gly Gln Leu
              110                 115                 120

CCT GGG GGC AAA GCA TCT TCA AAA GCA GGA TCT GCC CCC AGC TCC TTC      735
Pro Gly Gly Lys Ala Ser Ser Lys Ala Gly Ser Ala Pro Ser Ser Phe
          125                 130                 135

CTG CTG AAG AAG ACC AGG GAG CCT GGG ACC CCT CGA GAG CCC AAG GAG      783
Leu Leu Lys Lys Thr Arg Glu Pro Gly Thr Pro Arg Glu Pro Lys Glu
      140                 145                 150

CCG TTC CGC CCG CCC CCC ATC ACA CCC CAC GAA TAC ATG CTC TCC CTG      831
Pro Phe Arg Pro Pro Pro Ile Thr Pro His Glu Tyr Met Leu Ser Leu
  155                 160                 165                 170

TAC AGG ACG CTG TCC GAT GCT GAC AGA AAG GGA GGT AAC AGC AGC GTG      879
```

```
Tyr Arg Thr Leu Ser Asp Ala Asp Arg Lys Gly Gly Asn Ser Ser Val
                175                 180                 185

AAG TTG GAG GCT GGC CTG GCC AAC ACC ATC ACC AGC TTT ATT GAC AAA      927
Lys Leu Glu Ala Gly Leu Ala Asn Thr Ile Thr Ser Phe Ile Asp Lys
            190                 195                 200

GGG CAA GAT GAC CGA GGC CCT GCG GTC AGG AAG CAG AGG TAC GTG TTT      975
Gly Gln Asp Asp Arg Gly Pro Ala Val Arg Lys Gln Arg Tyr Val Phe
        205                 210                 215

GAC ATC AGT GCC TTG GAG AAG GAT GGG CTG TTG GGG GCT GAA CTG CGG     1023
Asp Ile Ser Ala Leu Glu Lys Asp Gly Leu Leu Gly Ala Glu Leu Arg
    220                 225                 230

ATC TTA CGG AAG AAG CCC TTG GAC GTG GCC AAG CCA GCG GTC CCC AGT     1071
Ile Leu Arg Lys Lys Pro Leu Asp Val Ala Lys Pro Ala Val Pro Ser
235                 240                 245                 250

AGC GGG CGG GTT GCC CAA CTG AAG CTG TCC AGC TGC CCC AGC GGC CGG     1119
Ser Gly Arg Val Ala Gln Leu Lys Leu Ser Ser Cys Pro Ser Gly Arg
                255                 260                 265

CAG CCG GCA GCC TTG CTG GAT GTG CGC TCC GTG CCA GGC CTG GAT GGA     1167
Gln Pro Ala Ala Leu Leu Asp Val Arg Ser Val Pro Gly Leu Asp Gly
            270                 275                 280

TCT GGC TGG GAG GTG TTC GAC ATC TGG AAG CTC TTC CGA AAT TTT AAG     1215
Ser Gly Trp Glu Val Phe Asp Ile Trp Lys Leu Phe Arg Asn Phe Lys
        285                 290                 295

AAC TCA GCG CAG CTG TGC CTG GAG CTG GAG GCC TGG GAA CGG GGC CGG     1263
Asn Ser Ala Gln Leu Cys Leu Glu Leu Glu Ala Trp Glu Arg Gly Arg
    300                 305                 310

GCC GTG GAC CTC CGT GGC CTG GGC TTT GAA CGC ACT GCC CGA CAG GTC     1311
Ala Val Asp Leu Arg Gly Leu Gly Phe Glu Arg Thr Ala Arg Gln Val
315                 320                 325                 330

CAC GAG AAA GCC TTG TTC CTA GTG TTT GGT CGT ACC AAG AAA CGG GAC     1359
His Glu Lys Ala L u Phe Leu Val Phe Gly Arg Thr Lys Lys Arg Asp
                335                 340                 345
```

```
CTG TTC TTT AAT GAG ATT AAG GCC CGC TCT GGC CAG GAT GAC AAG ACT      1407
Leu Phe Phe Asn Glu Ile Lys Ala Arg Ser Gly Gln Asp Asp Lys Thr
            350                 355                 360

GTG TAT GAA TAT TTG TTC AGC CAG CGG CGG AAA CGC CGG GCC CCA TTG      1455
Val Tyr Glu Tyr Leu Phe Ser Gln Arg Arg Lys Arg Arg Ala Pro Leu
        365                 370                 375

GCC AAT CGC CAG GGC AAG CGA CCC AGC AAG AAC CTC AAG GCT CGC TGC      1503
Ala Asn Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu Lys Ala Arg Cys
    380                 385                 390

AGT CGC AAG GCC TTG CAT GTC AAC TTC AAG GAC ATG GGC TGG GAC GAC      1551
Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly Trp Asp Asp
395                 400                 405                 410

TGG ATC ATC GCA CCT CTT GAG TAT GAG GCC TTC CAC TGC GAA GGA CTG      1599
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu
                415                 420                 425

TGT GAG TTC CCC TTG CGC TCC CAC TTG GAG CCC ACA AAC CAC GCA GTC      1647
Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val
            430                 435                 440

ATT CAG ACC CTA ATG AAC TCT ATG GAC CCT GAA TCC ACA CCA CCC ACT      1695
Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr
        445                 450                 455

TGT TGT GTG CCT ACA CGG CTG AGT CCT ATT AGC ATC CTC TTC ATC GAC      1743
Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp
    460                 465                 470

TCT GCC AAC AAC GTG GTG TAT AAA CAG TAC GAG GAC ATG GTC GTG GAA      1791
Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu
475                 480                 485                 490

TCT TGT GGC TGC AGG T AGCAGCACCG GCCCACCTGT CTTCCAGGGT GGCACATCCA    1847
Ser Cys Gly Cys Arg
            495

GAGACTACCC CCTCTACAGG TTCCTGGAGT AACAGAGAGC CTGTGAAGCT GCTGCCCGAA    1907

GTTTCCTGGC AGCCTGCAGG AAAGAGTTCT CAGCAGGCTT ACTCTCTGGA TGTGATCTGG    1967

ACTAAAGAGA TCACCTTCTG AAGATTCCTG CCCAAGGAAC AGACTCTGAG TGGGCCTGGG    2027
```

GCTCAGGAAA GGTGTTCTTA ATGAGATTCA GTTCACCATC TCTCCTGCCG GGGCCGGAGA 2087

CCTTCATTTC TCTCCAGACT CTCCAGAGAA GTTGTAGCTA TATCCTAAGC TCTTTAAGGG 2147

AGAGCTGTCT CCTCCTTGAA TCACCTTTGT GCCTGGTGAC TTTCTGCCAC GAGATGTTCA 2207

TTACAGGGGC TGGGCAAAGA AGGGGAAAGG GCTTGGGCAG GGGTGAAGAG AAGAGTATGA 2267

GCCTAATTAG ACTGTTAGAT TAAAATGTAC ATCGATGACA TAAAAGCTGA ATCTTCATGG 2327

CT 2329

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 495 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Arg Leu Pro Lys Leu Leu Thr Leu Leu Leu Trp His Leu Ala Trp
1 5 10 15

Leu Asp Leu Glu Leu Ile Cys Thr Val Leu Gly Ala Pro Asp Leu Gly
20 25 30

Gln Arg Thr Pro Gly Ala Lys Pro Gly Leu Thr Lys Ala Glu Ala Lys
35 40 45

Glu Arg Pro Pro Leu Ala Arg Asn Val Phe Arg Pro Gly Gly His Ile
50 55 60

Tyr Gly Val Gly Ala Thr Asn Ala Arg Ala Lys Gly Ser Ser Gly Gln
65 70 75 80

Thr Gln Ala Lys Lys Asp Glu Pro Arg Lys Met Pro Pro Arg Ser Gly
85 90 95

Gly Ser Glu Thr Lys Pro Gly Pro Ser Ser Gln Thr Arg Gln Ala Ala
100 105 110

Ala Arg Thr Val Thr Pro Lys Gly Gln Leu Pro Gly Gly Lys Ala Ser
115 120 125

Ser Lys Ala Gly Ser Ala Pro Ser Ser Phe Leu Leu Lys Lys Thr Arg
130 135 140

Glu Pro Gly Thr Pro Arg Glu Pro Lys Glu Pro Phe Arg Pro Pro Pro
145 150 155 160

Ile Thr Pro His Glu Tyr Met Leu Ser Leu Tyr Arg Thr Leu Ser Asp
165 170 175

Ala Asp Arg Lys Gly Gly Asn Ser Ser Val Lys Leu Glu Ala Gly Leu
180 185 190

Ala Asn Thr Ile Thr Ser Phe Ile Asp Lys Gly Gln Asp Asp Arg Gly
195 200 205

Pro Ala Val Arg Lys Gln Arg Tyr Val Phe Asp Ile Ser Ala Leu Glu
210 215 220

Lys Asp Gly Leu Leu Gly Ala Glu Leu Arg Ile Leu Arg Lys Lys Pro
225 230 235 240

Leu Asp Val Ala Lys Pro Ala Val Pro Ser Ser Gly Arg Val Ala Gln
245 250 255

Leu Lys Leu Ser Ser Cys Pro Ser Gly Arg Gln Pro Ala Ala Leu Leu
260 265 270

Asp Val Arg Ser Val Pro Gly Leu Asp Gly Ser Gly Trp Glu Val Phe
275 280 285

Asp Ile Trp Lys Leu Phe Arg Asn Phe Lys Asn Ser Ala Gln Leu Cys
290 295 300

Leu Glu Leu Glu Ala Trp Glu Arg Gly Arg Ala Val Asp Leu Arg Gly
305 310 315 320

Leu Gly Phe Glu Arg Thr Ala Arg Gln Val His Glu Lys Ala Leu Phe
325 330 335

Leu Val Phe Gly Arg Thr Lys Lys Arg Asp Leu Phe Phe Asn Glu Ile
340 345 350

```
Lys Ala Arg Ser Gly Gln Asp Asp Lys Thr Val Tyr Glu Tyr Leu Phe
        355                 360                 365

Ser Gln Arg Arg Lys Arg Arg Ala Pro Leu Ala Asn Arg Gln Gly Lys
    370                 375                 380

Arg Pro Ser Lys Asn Leu Lys Ala Arg Cys Ser Arg Lys Ala Leu His
385                 390                 395                 400

Val Asn Phe Lys Asp Met Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu
                405                 410                 415

Glu Tyr Glu Ala Phe His Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg
            420                 425                 430

Ser His Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu Met Asn
        435                 440                 445

Ser Met Asp Pro Glu Ser Thr Pro Pro Thr Cys Cys Val Pro Thr Arg
    450                 455                 460

Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val Val
465                 470                 475                 480

Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys Arg
                485                 490                 495
```

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(B) CLONE: GDF-1

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..124

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Arg Leu Arg Arg His Thr Glu Pro Arg Val Glu Val Gly Pro Val Gly
1 5 10 15

Thr Cys Arg Thr Arg Arg Leu His Val Ser Phe Arg Glu Val Gly Trp
20 25 30

His Arg Trp Val Ile Ala Pro Arg Gly Phe Leu Ala Asn Phe Cys Gln
35 40 45

Gly Thr Cys Ala Leu Pro Glu Thr Leu Arg Gly Pro Gly Gly Pro Pro
50 55 60

Ala Leu Asn His Ala Val Leu Arg Ala Leu Met His Ala Ala Ala Pro
65 70 75 80

Thr Pro Gly Ala Gly Ser Pro Cys Cys Val Pro Glu Arg Leu Ser Pro
85 90 95

Ile Ser Val Leu Phe Phe Asp Asn Glu Asp Asn Val Val Leu Arg His
100 105 110

Tyr Glu Asp Met Val Val Asp Glu Cys Gly Cys Arg
115 120

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: GDF-3

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..118

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Arg Lys Arg Arg Ala Ala Ile Ser Val Pro Lys Gly Phe Cys Arg Asn
1 5 10 15

Phe Cys His Arg His Gln Leu Phe Ile Asn Phe Gln Asp Leu Gly Trp
20 25 30

His Lys Trp Val Ile Ala Pro Lys Gly Phe Met Ala Asn Tyr Cys His
35 40 45

Gly Glu Cys Pro Phe Ser Met Thr Thr Tyr Leu Asn Ser Ser Asn Tyr
50 55 60

Ala Phe Met Gln Ala Leu Met His Met Ala Asp Pro Lys Val Pro Lys
65 70 75 80

Ala Val Cys Val Pro Thr Lys Leu Ser Pro Ile Ser Met Leu Tyr Gln
85 90 95

Asp Ser Asp Lys Asn Val Ile Leu Arg His Tyr Glu Asp Met Val Val
100 105 110

Asp Glu Cys Gly Cys Gly
115

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: GDF-5

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Pro Leu Ala Asn Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu Lys Ala
1 5 10 15

Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly Trp
20 25 30

Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu
35 40 45

Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His
50 55 60

Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro
65 70 75 80

Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe
85 90 95

Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val
100 105 110

Val Glu Ser Cys Gly Cys Arg
115

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: GDF-9

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Ser Phe Asn Leu Ser Glu Tyr Phe Lys Gln Phe Leu Phe Pro Gln Asn
1 5 10 15

Glu Cys Glu Leu His Asp Phe Arg Leu Ser Phe Ser Gln Leu Lys Trp
20 25 30

Asp Asn Trp Ile Val Ala Pro His Arg Tyr Asn Pro Arg Tyr Cys Lys
35 40 45

Gly Asp Cys Pro Arg Ala Val Arg His Arg Tyr Gly Ser Pro Val His
50 55 60

Thr Met Val Gln Asn Ile Ile Tyr Glu Lys Leu Asp Pro Ser Val Pro
65 70 75 80

Arg Pro Ser Cys Val Pro Gly Lys Tyr Ser Pro Leu Ser Val Leu Thr
85 90 95

Ile Glu Pro Asp Gly Ser Ile Ala Tyr Lys Glu Tyr Glu Asp Met Ile
100 105 110

Ala Thr Arg Cys Thr Cys Arg
115

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: BMP-2

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..118

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Arg Glu Lys Arg Gln Ala Lys His Lys Gln Arg Lys Arg Leu Lys Ser
1 5 10 15

Ser Cys Lys Arg His Pro Leu Tyr Val Asp Phe Ser Asp Val Gly Trp
20 25 30

Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr His Ala Phe Tyr Cys His
35 40 45

Gly Glu Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His
50 55 60

Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Lys Ile Pro Lys
65 70 75 80

Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu
85 90 95

Asp Glu Asn Glu Lys Val Val Leu Lys Asn Tyr Gln Asp Met Val Val
100 105 110

Glu Gly Cys Gly Cys Arg
115

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: BMP-4

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..118

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Lys Arg Ser Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys
1 5 10 15

Asn Cys Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp
20 25 30

Asn Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His
35 40 45

Gly Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn His
50 55 60

Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile Pro Lys
65 70 75 80

Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met Leu Tyr Leu
85 90 95

Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln Glu Met Val Val
100 105 110

Glu Gly Cys Gly Cys Arg
115

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: Vgr-1

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ser Arg Gly Ser Gly Ser Ser Asp Tyr Asn Gly Ser Glu Leu Lys Thr
1 5 10 15

Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Gln Asp Leu Gly Trp
20 25 30

Gln Asp Trp Ile Ile Ala Pro Lys Gly Tyr Ala Ala Asn Tyr Cys Asp
35 40 45

Gly Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His
50 55 60

Ala Ile Val Gln Thr Leu Val His Leu Met Asn Pro Glu Tyr Val Pro
65 70 75 80

Lys Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr
85 90 95

Phe Asp Asp Asn Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val
100 105 110

Val Arg Ala Cys Gly Cys His
115

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: OP-1

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Leu Arg Met Ala Asn Val Ala Glu Asn Ser Ser Ser Asp Gln Arg Gln
1 5 10 15

Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp
20 25 30

Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Tyr Tyr Cys Glu
35 40 45

Gly Glu Cys Ala Phe Pro Leu Asn Ser Tyr Met Asn Ala Thr Asn His
50 55 60

Ala Ile Val Gln Thr Leu Val His Phe Ile Asn Pro Glu Thr Val Pro
65 70 75 80

Lys Pro Cys Cys Ala Pro Thr Gln Leu Asn Ala Ile Ser Val Leu Tyr
85 90 95

Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val
100 105 110

Val Arg Ala Cys Gly Cys His
115

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: BMP-5

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..119

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Ser Arg Met Ser Ser Val Gly Asp Tyr Asn Thr Ser Glu Gln Lys Gln
1 5 10 15

Ala Cys Lys Lys His Glu Leu Tyr Val Ser Phe Arg Asp Leu Gly Trp
20 25 30

Gln Asp Trp Ile Ile Ala Pro Glu Gly Tyr Ala Ala Phe Tyr Cys Asp
35 40 45

Gly Glu Cys Ser Phe Pro Leu Asn Ala His Met Asn Ala Thr Asn His
50 55 60

Ala Ile Val Gln Thr Leu Val His Leu Met Phe Pro Asp His Val Pro
65 70 75 80

Lys Pro Cys Cys Ala Pro Thr Lys Leu Asn Ala Ile Ser Val Leu Tyr
85 90 95

Phe Asp Asp Ser Ser Asn Val Ile Leu Lys Lys Tyr Arg Asn Met Val
100 105 110

Val Arg Ser Cys Gly Cys His
115

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: BMP-3

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..120

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Glu Gln Thr Leu Lys Lys Ala Arg Arg Lys Gln Trp Ile Glu Pro Arg
1 5 10 15

Asn Cys Ala Arg Arg Tyr Leu Lys Val Asp Phe Ala Asp Ile Gly Trp
20 25 30

Ser Glu Trp Ile Ile Ser Pro Lys Ser Phe Asp Ala Tyr Tyr Cys Ser
35 40 45

Gly Ala Cys Gln Phe Pro Met Pro Lys Ser Leu Lys Pro Ser Asn His
50 55 60

Ala Thr Ile Gln Ser Ile Val Arg Ala Val Gly Val Val Pro Gly Ile
65 70 75 80

Pro Glu Pro Cys Cys Val Pro Glu Lys Met Ser Ser Leu Ser Ile Leu
85 90 95

Phe Phe Asp Glu Asn Lys Asn Val Val Leu Lys Val Tyr Pro Asn Met
100 105 110

Thr Val Glu Ser Cys Ala Cys Arg
115 120

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: MIS

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..116

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Gly Pro Gly Arg Ala Gln Arg Ser Ala Gly Ala Thr Ala Ala Asp Gly
1 5 10 15

Pro Cys Ala Leu Arg Glu Leu Ser Val Asp Leu Arg Ala Glu Arg Ser
20 25 30

Val Leu Ile Pro Glu Thr Tyr Gln Ala Asn Asn Cys Gln Gly Val Cys
35 40 45

Gly Trp Pro Gln Ser Asp Arg Asn Pro Arg Tyr Gly Asn His Val Val
50 55 60

Leu Leu Leu Lys Met Gln Ala Arg Gly Ala Ala Leu Ala Arg Pro Pro
65 70 75 80

Cys Cys Val Pro Thr Ala Tyr Ala Gly Lys Leu Leu Ile Ser Leu Ser
85 90 95

Glu Glu Arg Ile Ser Ala His His Val Pro Asn Met Val Ala Thr Glu
100 105 110

Cys Gly Cys Arg
115

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: Inhibit-alpha

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..122

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Ala Leu Arg Leu Leu Gln Arg Pro Pro Glu Glu Pro Ala Ala His Ala
1 5 10 15

Asn Cys His Arg Val Ala Leu Asn Ile Ser Phe Gln Glu Leu Gly Trp
20 25 30

Glu Arg Trp Ile Val Tyr Pro Pro Ser Phe Ile Phe His Tyr Cys His
35 40 45

Gly Gly Cys Gly Leu His Ile Pro Pro Asn Leu Ser Leu Pro Val Pro
50 55 60

Gly Ala Pro Pro Thr Pro Ala Gln Pro Tyr Ser Leu Leu Pro Gly Ala
65 70 75 80

Gln Pro Cys Cys Ala Ala Leu Pro Gly Thr Met Arg Pro Leu His Val
85 90 95

Arg Thr Thr Ser Asp Gly Gly Tyr Ser Phe Lys Tyr Glu Thr Val Pro
100 105 110

Asn Leu Leu Thr Gln His Cys Ala Cys Ile
115 120

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 122 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: Inhibin-beta-alpha

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..122

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

His Arg Arg Arg Arg Arg Gly Leu Glu Cys Asp Gly Lys Val Asn Ile
1 5 10 15

Cys Cys Lys Lys Gln Phe Phe Val Ser Phe Lys Asp Ile Gly Trp Asn
20 25 30

Asp Trp Ile Ile Ala Pro Ser Gly Tyr His Ala Asn Tyr Cys Glu Gly
35 40 45

Glu Cys Pro Ser His Ile Ala Gly Thr Ser Gly Ser Ser Leu Ser Phe
50 55 60

His Ser Thr Val Ile Asn His Tyr Arg Met Arg Gly His Ser Pro Phe
65 70 75 80

Ala Asn Leu Lys Ser Cys Cys Val Pro Thr Lys Leu Arg Pro Met Ser
85 90 95

Met Leu Tyr Tyr Asp Asp Gly Gln Asn Ile Ile Lys Lys Asp Ile Gln
100 105 110

Asn Met Ile Val Glu Glu Cys Gly Cys Ser
115 120

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 121 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: Inhibin-beta-beta

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..121

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

```
His Arg Ile Arg Lys Arg Gly Leu Glu Cys Asp Gly Arg Thr Asn Leu
1               5                  10                  15

Cys Cys Arg Gln Gln Phe Phe Ile Asp Phe Arg Leu Ile Gly Trp Asn
            20                  25                  30

Asp Trp Ile Ile Ala Pro Thr Gly Tyr Tyr Gly Asn Tyr Cys Glu Gly
        35                  40                  45

Ser Cys Pro Ala Tyr Leu Ala Gly Val Pro Gly Ser Ala Ser Ser Phe
    50                  55                  60

His Thr Ala Val Val Asn Gln Tyr Arg Met Arg Gly Leu Asn Pro Gly
65                  70                  75                  80

Thr Val Asn Ser Cys Cys Ile Pro Thr Lys Leu Ser Thr Met Ser Met
                85                  90                  95

Leu Tyr Phe Asp Asp Glu Tyr Asn Ile Val Lys Arg Asp Val Pro Asn
            100                 105                 110

Met Ile Val Glu Glu Cys Gly Cys Ala
        115                 120
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 115 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
- (B) CLONE: TGF-beta-1

(ix) FEATURE:
- (A) NAME/KEY: Protein
- (B) LOCATION: 1..115

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

His Arg Arg Ala Leu Asp Thr Asn Tyr Cys Phe Ser Ser Thr Glu Lys
1 5 10 15

Asn Cys Cys Val Arg Gln Leu Tyr Ile Asp Phe Arg Lys Asp Leu Gly
20 25 30

Trp Lys Trp Ile His Glu Pro Lys Gly Tyr His Ala Asn Phe Cys Leu
35 40 45

Gly Pro Cys Pro Tyr Ile Trp Ser Leu Asp Thr Gln Tyr Ser Lys Val
50 55 60

Leu Ala Leu Tyr Asn Gln His Asn Pro Gly Ala Ser Ala Ala Pro Cys
65 70 75 80

Cys Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val Tyr Tyr Val Gly
85 90 95

Arg Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile Val Arg Ser Cys
100 105 110

Lys Cys Ser
115

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 115 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: TGF-beta-2

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..115

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Lys Lys Arg Ala Leu Asp Ala Ala Tyr Cys Phe Arg Asn Val Gln Asp
1 5 10 15

Asn Cys Cys Leu Arg Pro Leu Tyr Ile Asp Phe Lys Arg Asp Leu Gly
20 25 30

Trp Lys Trp Ile His Glu Pro Lys Gly Tyr Asn Ala Asn Phe Cys Ala
35 40 45

Gly Ala Cys Pro Tyr Leu Trp Ser Ser Asp Thr Gln His Ser Arg Val
50 55 60

Leu Ser Leu Tyr Asn Thr Ile Asn Pro Glu Ala Ser Ala Ser Pro Cys
65 70 75 80

Cys Val Ser Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Ile Gly
85 90 95

Lys Thr Pro Lys Ile Glu Gln Leu Ser Asn Met Ile Val Lys Ser Cys
100 105 110

Lys Cys Ser
115

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 115 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vii) IMMEDIATE SOURCE:
(B) CLONE: TGF-beta-3

(ix) FEATURE:
(A) NAME/KEY: Protein
(B) LOCATION: 1..115

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Lys Lys Arg Ala Leu Asp Thr Asn Tyr Cys Phe Arg Asn Leu Glu Glu
1 5 10 15

Asn Cys Cys Val Arg Pro Leu Tyr Ile Asp Phe Arg Gln Asp Leu Gly
20 25 30

Trp Lys Trp Val His Glu Pro Lys Gly Tyr Tyr Ala Asn Phe Cys Ser
35 40 45

Gly Pro Cys Pro Tyr Leu Arg Ser Ala Asp Thr Thr His Ser Thr Val
50 55 60

Leu Gly Leu Tyr Asn Thr Leu Asn Pro Glu Ala Ser Ala Ser Pro Cys
65 70 75 80

Cys Val Pro Gln Asp Leu Glu Pro Leu Thr Ile Leu Tyr Tyr Val Gly
85 90 95

Arg Thr Pro Lys Val Glu Gln Leu Ser Asn Met Val Val Lys Ser Cys
100 105 110

Lys Cys Ser
115

Although the invention has been described with reference to the presently preferred embodiment, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

CLAIMS

1. Substantially pure growth differentiation factor-5 (GDF-5) and functional fragments thereof.

2. An isolated polynucleotide sequence encoding the GDF-5 polypeptide of claim 1.

3. The polynucleotide sequence of claim 2, wherein the polynucleotide is isolated from a mammalian cell.

4. The polynucleotide of claim 3, wherein the mammalian cell is selected from the group consisting of mouse, rat, and human cell.

5. An expression vector including the polynucleotide of claim 2.

6. The vector of claim 5, wherein the vector is a plasmid.

7. The vector of claim 5, wherein the vector is a virus.

8. A host cell stably transformed with the vector of claim 5.

9. The host cell of claim 8, wherein the cell is prokaryotic.

10. The host cell of claim 8, wherein the cell is eukaryotic.

11. Antibodies reactive with the polypeptide of claim 1 or fragments thereof.

12. The antibodies of claim 11, wherein the antibodies are polyclonal.

13. The antibodies of claim 11, wherein the antibodies are monoclonal.

14. A method of detecting a cell proliferative disorder comprising contacting the antibody of claim 11 with a specimen of a subject suspected of having a GDF-5 associated disorder and detecting binding of the antibody.

15. The method of claim 14, wherein the cell proliferative disorder is a uterine neoplasm or endometriosis.

16. The method of claim 14, wherein the cell proliferative disorder is a skeletal disorder.

17. The method of claim 14, wherein the detecting is *in vivo.*

18. The method of claim 17, wherein the antibody is detectably labeled.

19. The method of claim 18, wherein the detectable label is selected from the group consisting of a radioisotope, a fluorescent compound, a bioluminescent compound and a chemiluminescent compound.

20. The method of claim 14, wherein the detection is *in vitro.*

21. The method of claim 20, wherein the antibody is detectably labeled.

22. The method of claim 21, wherein the label is selected from the group consisting of a radioisotope, a fluorescent compound, a bioluminescent compound, a chemoluminescent compound and an enzyme.

23. A method of treating a cell proliferative disorder associated with expression of GDF-5, comprising contacting the cells with a reagent which suppresses the GDF-5 activity.

24. The method of claim 23, wherein the reagent is an anti-GDF-5 antibody.

25. The method of claim 23, wherein the reagent is a GDF-5 antisense sequence.

26. The method of claim 23, wherein the cell proliferative disorder is a uterine neoplasm or endometriosis.

27. The method of claim 23, wherein the cell proliferative disorder is a skeletal disorder.

28. The method of claim 23, wherein the reagent which suppresses GDF-5 activity is introduced to a cell using a vector.

29. The method of claim 28, wherein the vector is a colloidal dispersion system.

30. The method of claim 29, wherein the colloidal dispersion system is a liposome.

31. The method of claim 30, wherein the liposome is essentially target specific.

32. The method of claim 31, wherein the liposome is anatomically targeted.

33. The method of claim 31, wherein the liposome is mechanistically targeted.

34. The method of claim 33, wherein the mechanistic targeting is passive.

35. The method of claim 33, wherein the mechanistic targeting is active.

36. The method of claim 35, wherein the liposome is actively targeted by coupling with a moiety selected from the group consisting of a sugar, a glycolipid, and a protein.

37. The method of claim 36, wherein the protein moiety is an antibody.

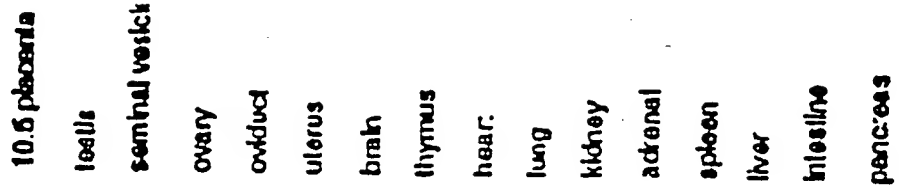
38. The method of claim 37, wherein the vector is a virus.

39. The method of claim 38, wherein the virus is an RNA virus.

40. The method of claim 39, wherein the RNA virus is a retrovirus.

41. The method of claim 40, wherein the retrovirus is essentially target specific.

a


10.5 placenta
testis
seminal vesicle
ovary
oviduct
uterus
brain
thymus
heart
lung
kidney
adrenal
spleen
liver
intestine
pancreas

b

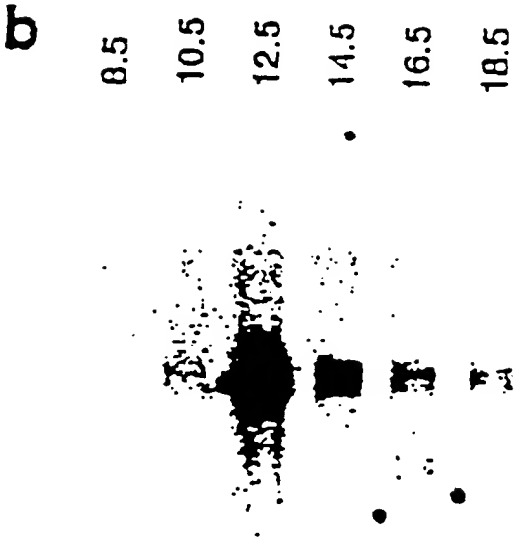

8.5
10.5
12.5
14.5
16.5
18.5

FIGURE 1

```
   1  TTCAAGCCCTCAGTCAGTTGTGCGGGAGAAAGGGGGCGGTCGGCTTTCTCCTTTCAAGAA   60
  61  CGAGTTATTTTCAGCTGCTGACTGGAGACGGTGCACGTCTGGACACGGGAGCACTTCCAC  120
 121  TATGGGACTGGATACAGACACACGCCCGGCGGACTTCAAGACACTCAGACTGAGGAGAAA  180
 181  GCCCTGCCTGCTGCTGCTGCTGCTGCTGCCACCGCTGCCTCTGAAGACCCACTCCTT    240
 241  TCATGGTTTTTCCTGCCAAGCCAGAGGCACCTTCGCTGCTACGGCCTTTCTCTGTGGTGT  300
 301  CATTCAGCGGCTGGCCAGAGGATGAGACTCCCCAAACTCCTCACTCTTTTGCTGTGGCAC  360
                    M  R  L  P  K  L  L  T  L  L  L  W  H
 361  CTGGCTTGGCTGGACCTGGAACTCATCTGCACTGTGCTGGGTGCCCCTGACTTAGGACAG  420
      L  A  W  L  D  L  E  L  I  C  T  V  L  G  A  P  D  L  G  Q
 421  AGAACCCCAGGGGCCAAGCCAGGGTTGACCAAAGCGGAGGCCAAGGAGAGGCCACCCCTG  480
      R  T  P  G  A  K  P  G  L  T  K  A  E  A  K  E  R  P  P  L
 481  GCCAGGAATGTCTTTAGGCCAGGGGGTCATATCTATGGTGTGGGGGCCACCAATGCCAGG  540
      A  R  N  V  F  R  P  G  G  H  I  Y  G  V  G  A  T  N  A  R
 541  GCCAAGGGAAGCTCTGGGCAGACACAGGCCAAGAAGGATGAACCCAGAAAGATGCCCCCC  600
      A  K  G  S  S  G  Q  T  Q  A  K  K  D  E  P  R  K  M  P  P
 601  AGATCCGGTGGCTCTGAAACCAAGCCAGGACCCTCTTCCCAGACTAGACAGGCTGCAGCC  660
      R  S  G  G  S  E  T  K  P  G  P  S  S  Q  T  R  Q  A  A  A
 661  CGGACTGTAACCCCAAAAGGACAGCTTCCTGGGGGCAAAGCATCTTCAAAAGCAGGATCT  720
      R  T  V  T  P  K  G  Q  L  P  G  G  K  A  S  S  K  A  G  S
 721  GCCCCCAGCTCCTTCCTGCTGAAGAAGACCAGGGAGCCTGGGACCCCTCGAGAGCCCAAG  780
      A  P  S  S  F  L  L  K  K  T  R  E  P  G  T  P  R  E  P  K
 781  GAGCCGTTCCGCCCGCCCCCCATCACACCCCACGAATACATGCTCTCCCTGTACAGGACG  840
      E  P  F  R  P  P  P  I  T  P  H  E  Y  M  L  S  L  Y  R  T
 841  CTGTCCGATGCTGACAGAAAGGGAGGTAACAGCAGCGTGAAGTTGGAGGCTGGCCTGGCC  900
      L  S  D  A  D  R  K  G  G  [N  S  S]  V  K  L  E  A  G  L  A
 901  AACACCATCACCAGCTTTATTGACAAAGGGCAAGATGACCGAGGCCCTGCGGTCAGGAAG  960
      N  T  I  T  S  F  I  D  K  G  Q  D  D  R  G  P  A  V  R  K
 961  CAGAGGTACGTGTTTGACATCAGTGCCTTGGAGAAGGATGGGCTGTTGGGGGCTGAACTG 1020
      Q  R  Y  V  F  D  I  S  A  L  E  K  D  G  L  L  G  A  E  L
1021  CGGATCTTACGGAAGAAGCCCTTGGACGTGGCCAAGCCAGCGGTCCCCAGTAGCGGGCGG 1080
      R  I  L  R  K  K  P  L  D  V  A  K  P  A  V  P  S  S  G  R
1081  GTTGCCCAACTGAAGCTGTCCAGCTGCCCCAGCGGCCGGCAGCCGGCAGCCTTGCTGGAT 1140
      V  A  Q  L  K  L  S  S  C  P  S  G  R  Q  P  A  A  L  L  D
1141  GTGCGCTCCGTGCCAGGCCTGGATGGATCTGGCTGGGAGGTGTTCGACATCTGGAAGCTC 1200
      V  R  S  V  P  G  L  D  G  S  G  W  E  V  F  D  I  W  K  L
1201  TTCCGAAATTTTAAGAACTCAGCGCAGCTGTGCCTGGAGCTGGAGGCCTGGGAACGGGGC 1260
      F  R  N  F  K  N  S  A  Q  L  C  L  E  L  E  A  W  E  R  G
1261  CGGGCCGTGGACCTCCGTGGCCTGGGCTTTGAACGCACTGCCCGACAGGTCCACGAGAAA 1320
      R  A  V  D  L  R  G  L  G  F  E  R  T  A  R  Q  V  H  E  K
1321  GCCTTGTTCCTAGTGTTTGGTCGTACCAAGAAACGGGACCTGTTCTTTAATGAGATTAAG 1380
      A  L  F  L  V  F  G  R  T  K  K  R  D  L  F  F  N  E  I  K
1381  GCCCGCTCTGGCCAGGATGACAAGACTGTGTATGAATATTTGTTCAGCCAGCGGCGGAAA 1440
      A  R  S  G  Q  D  D  K  T  V  Y  E  Y  L  F  S  Q  [R  R  K]
1441  CGCCGGGCCCCATTGGCCAATCGCCAGGGCAAGCGACCCAGCAAGAACCTCAAGGCTCGC 1500
      [R  R]  A  P  L  A  N  R  Q  G  K  R  P  S  K  N  L  K  A  R
1501  TGCAGTCGCAAGGCCTTGCATGTCAACTTCAAGGACATGGGCTGGGACGACTGGATCATC 1560
      C  S  R  K  A  L  H  V  N  F  K  D  M  G  W  D  D  W  I  I
1561  GCACCTCTTGAGTATGAGGCCTTCCACTGCGAAGGACTGTGTGAGTTCCCCTTGCGCTCC 1620
      A  P  L  E  Y  E  A  F  H  C  E  G  L  C  E  F  P  L  R  S
1621  CACTTGGAGCCCACAAACCACGCAGTCATTCAGACCCTAATGAACTCTATGGACCCTGAA 1680
      H  L  E  P  T  N  H  A  V  I  Q  T  L  M  N  S  M  D  P  E
1681  TCCACACCACCCACTTGTTGTGTGCCTACACGGCTGAGTCCTATTAGCATCCTCTTCATC 1740
      S  T  P  P  T  C  C  V  P  T  R  L  S  P  I  S  I  L  F  I
1741  GACTCTGCCAACAACGTGGTGTATAAACAGTACGAGGACATGGTCGTGGAATCTTGTGGC 1800
      D  S  A  N  N  V  V  Y  K  Q  Y  E  D  M  V  V  E  S  C  G
1801  TGCAGGTAGCAGCACCGGCCCACCTGTCTTCCAGGGTGGCACATCCAGAGACTACCCCCT 1860
      C  R  *
1861  CTACAGGTTCCTGGAGTAACAGAGAGCCTGTGAAGCTGCTGCCCGAAGTTTCCTGGCAGC 1920
1921  CTGCAGGAAAGAGTTCTCAGCAGGCTTACTCTCTGGATGTGATCTGGACTAAAGAGATCA 1980
1981  CCTTCTGAAGATTCCTGCCCAAGGAACAGACTCTGAGTGGGCCTGGGGCTCAGGAAAGGT 2040
2041  GTTCTTAATGAGATTCAGTTCACCATCTCTCCTGCCGGGGCCGGAGACCTTCATTTCTCT 2100
2101  CCAGACTCTCCAGAGAAGTTGTAGCTATATCCTAAGCTCTTTAAGGGAGAGCTGTCTCCT 2160
2161  CCTTGAATCACCTTTGTGCCTGGTGACTTTCTGCCACGAGATGTTCATTACAGGGGCTGG 2220
2221  GCAAAGAAGGGGAAAGGGCTTGGGCAGGGGTGAAGAGAAGAGTATGAGCCTAATTAGACT 2280
2281  GTTAGATTAAAATGTACATCGATGACATAAAAGCTGAATCTTCATGGCT  2329
```

FIGURE 2

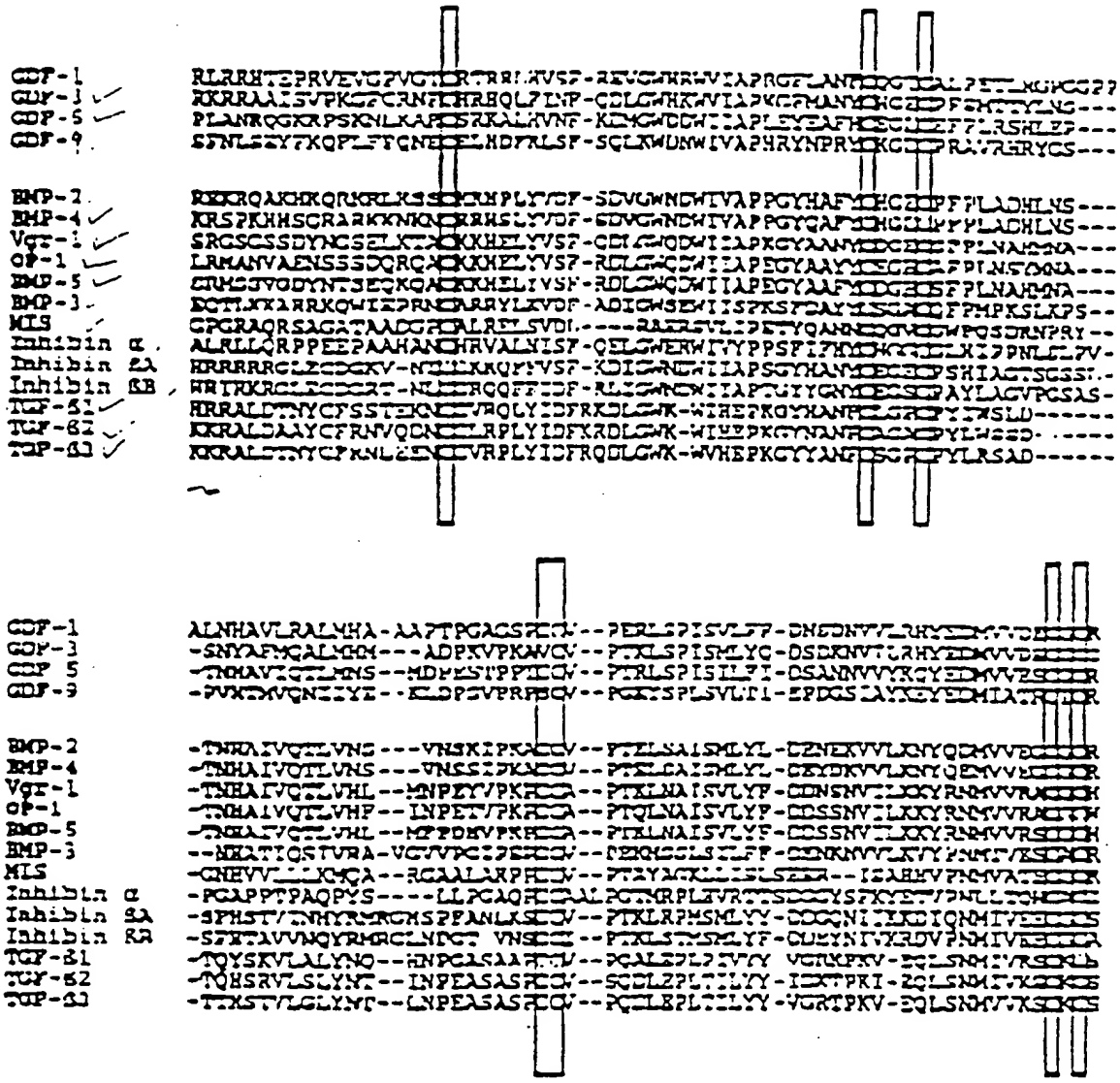


FIGURE 3a

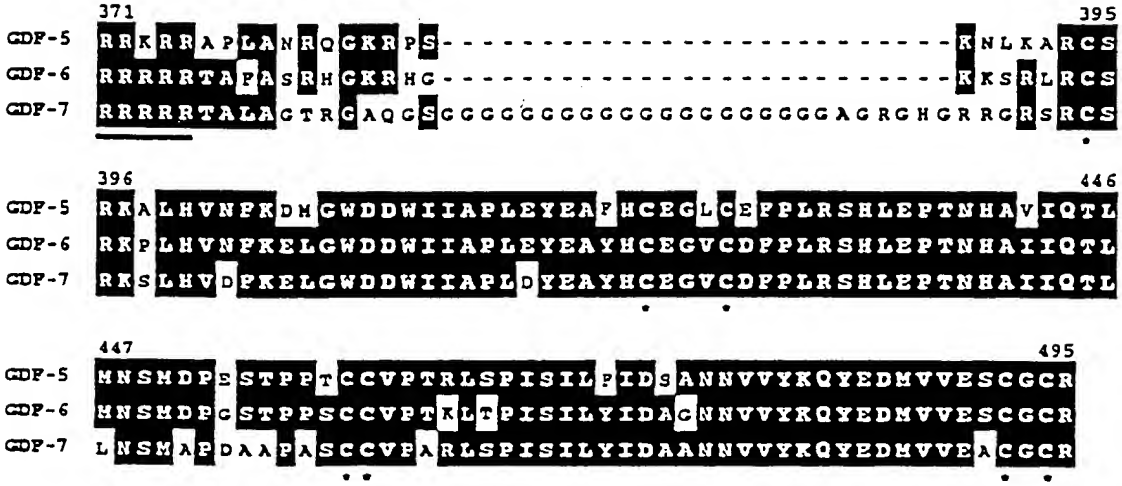

371
395
GDF-5 RRKRRAPLANRQGKRPS----------KNLKARCS
GDF-6 RRRRRTAPASRHGKRHG----------KKSRLRCS
GDF-7 RRRRRTALAGTRGAQGSGGGGGGGGGGGGGGGGGGAGRGHGRRGRSRCS
396
446
GDF-5 RKALHVNFKDMGWDDWIIAPLEYEAFHCEGLCEFPLRSHLEPTNHAVIQTL
GDF-6 RKPLHVNFKELGWDDWIIAPLEYEAYHCEGVCDFPLRSHLEPTNHAIIQTL
GDF-7 RKSLHVDPKELGWDDWIIAPLDYEAYHCEGVCDFPLRSHLEPTNHAIIQTL
447
495
GDF-5 MNSMDPESTPPTCCVPTRLSPISILFIDSANNVVYKQYEDMVVESCGCR
GDF-6 MNSMDPGSTPPSCCVPTKLTPISILYIDAGNNVVYKQYEDMVVESCGCR
GDF-7 LNSMAPDAAPASCCVPARLSPISILYIDAANNVVYKQYEDMVVEACGCR

FIGURE 3b

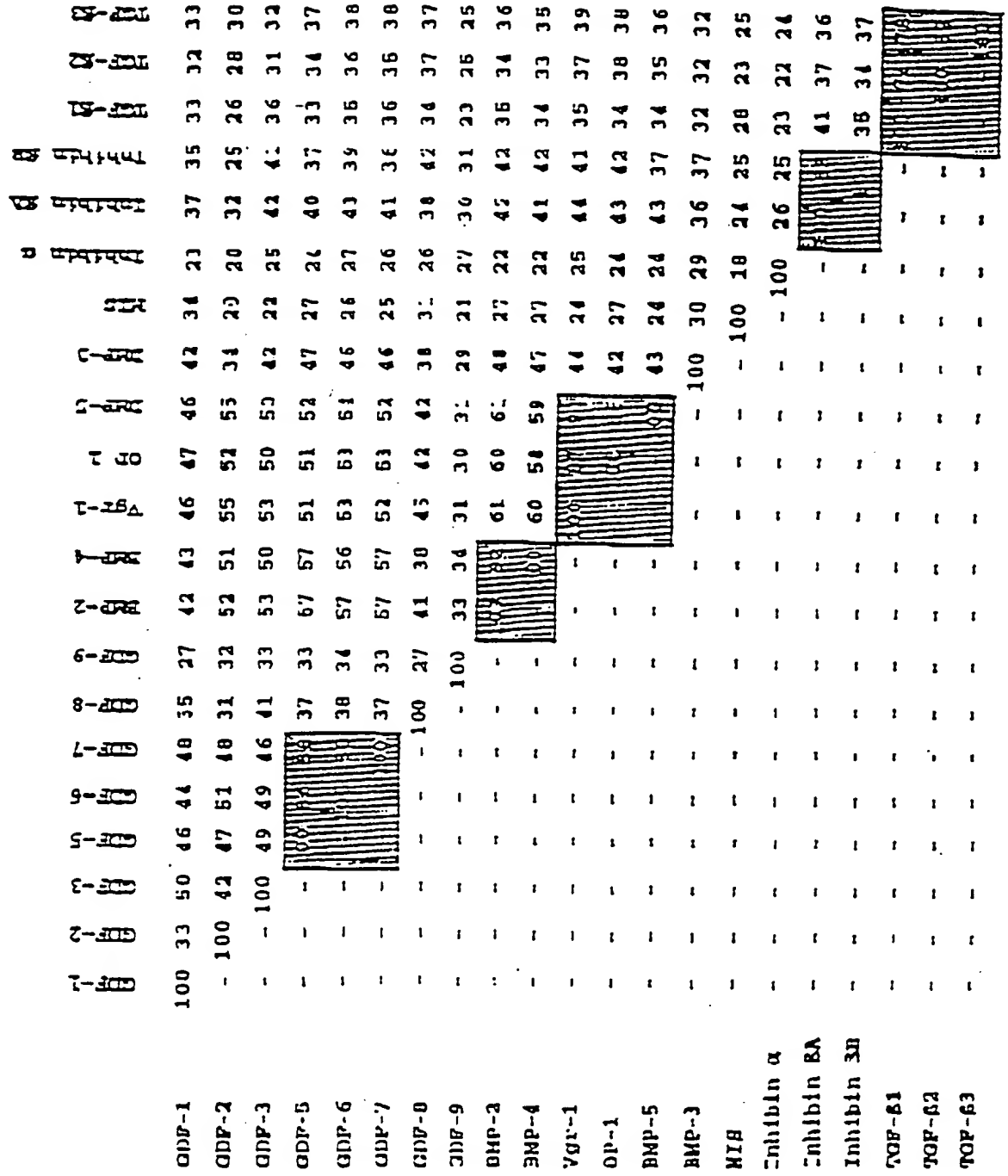


FIGURE 4

FIGURE 5a

FIGURE 5b

FIGURE 5c

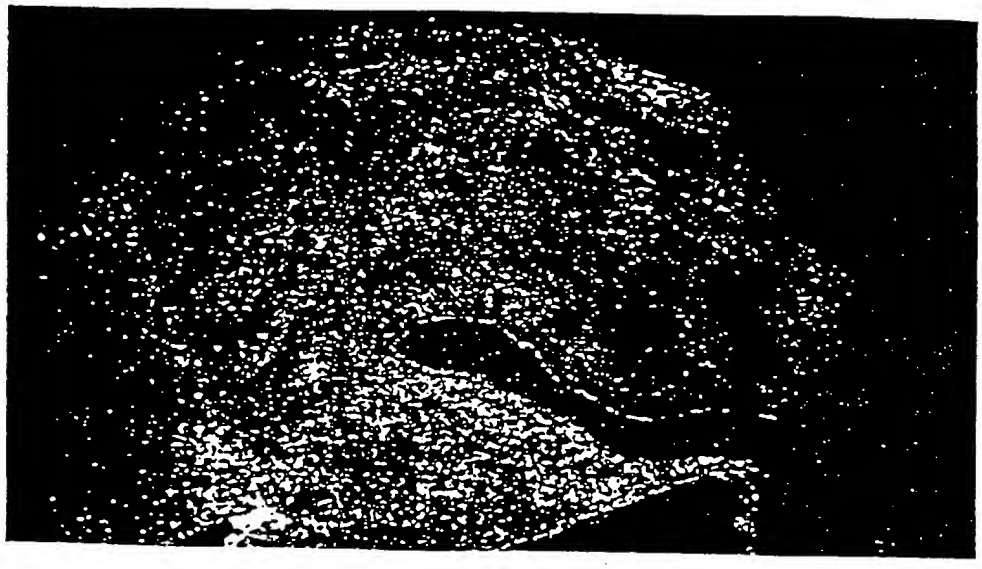


FIGURE 5d

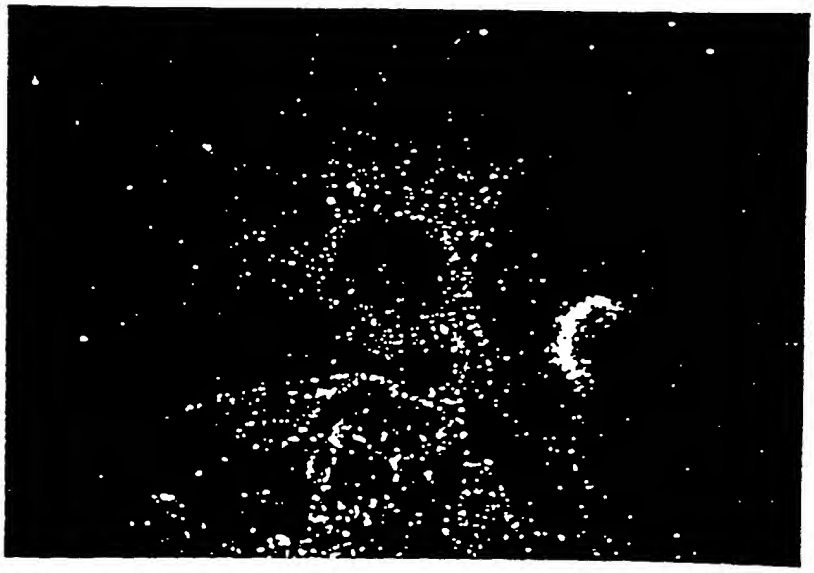


FIGURE 5e

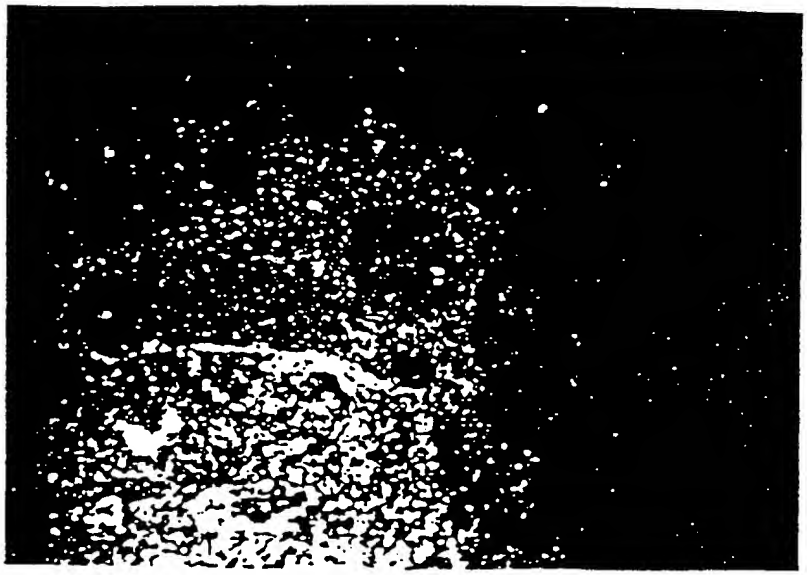


FIGURE 5f

INTERNATIONAL SEARCH REPORT

Int. ational application No.
PCT/US94/00657

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : Please See Extra Sheet.

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X _ Y	Molecular Endocrinology, Vol. 4, issued 1990, Lee, S.-J., "Identification Of A Novel Member (GDF-1) Of The Transforming Growth Factor-ß Superfamily", pages 1034-1040, see entire document.	2-6, 8, 9 _____ 1-13
Y	Proc. Natl. Acad. Sci., USA, Vol. 88, Issued May 1991, Lee, S.-J., "Expression Of Growth/Differentiation Factor 1 In The Nervous System: Conservation Of A Bicistronic Structure", pages 4250-4254, see entire document.	1-13
Y	Meth. Enzymol., Vol. 100, issued 1983, Beltz et al.., ""Isolation of Multigene Families And Determination Of Homologies By Filter Hybridization Methods", pages 266-285, see entire document.	1-13

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be part of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 APRIL 1994	APR 15 1994
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. NOT APPLICABLE	Authorized officer CHRISTOPHER S. LOW Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

Inte. .ional application No.
PCT/US94/00657

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Sambrook et al., "Molecular Cloning, A Laboratory Manual", Second Edition, published 1989 by Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pages 11.2-11.11, 11.17-11.19, see all cited pages.	1-13
Y	US, A, 4,675,285 (CLARK et al.) 23 June 1987, see entire document.	1-13

INTERNATIONAL SEARCH REPORT

Int tional application No.
PCT/US94/00657

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
1-13

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark n Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Int. ional application No.
PCT/US94/00657

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (5):

C07K 3/00, 13/00, 15/28, 17/00; C07H 15/12, 17/00; C12N 15/70, 15/79, 15/00, 7/00, 1/20, 5/16, 1/21

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

530 / 350, 399, 387.1; 536 / 23.1, 23.51; 435 / 320.1, 235.1, 252.3, 240.2, 172.3

B. FIELDS SEARCHED
Minimum documentation searched
Classification System: U.S.

530 / 350, 399, 387.1; 536 / 23.1, 23.51; 435 / 320.1, 235.1, 252.3, 240.2, 172.3

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

Automated Patent System - USPAT, JPOABS
Dialog one search files: 5, 265, 266
EMBL-NEW, GenBank 80, GenBank-NEW, N-GeneSeq 13, UEMBL 37_80, A-GeneSeq 13, PIR 38, and Swiss-PRot 27

Search terms: growth, differentiation, factor 5, vector, viral, virus, plasmid

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This international application contains the following inventions or groups of inventions which are not so linked by the identical special technical feature so as to form a single inventive concept under the criteria of PCT Rule 13.2.

I.Claims 1, drawn to growth differentiation factor5 (GDF-5) are for example, classified in Class 530, subclasses 350 and 399.

II.Claims 2-10, drawn to polynucleotides encoding growth differentiation factor5 as well as vectors and host cells containing same are for example, classified in Class 536, subclass 23.1 and 23.51 and Class 435 subclasses 320.1, 235.1, 252.3, and 240.2.

III.Claims 11-13, drawn to antibodies to growth differentiation factor-5 are for example, classified in at least Class 530, subclass 387.1.

IV.Claims 14-22, drawn to a method of detecting a cell proliferative disorder by binding an antibody to the sample specimen are for example, classified in Class 435, subclass 7.1.

V.Claims 23-41, drawn to a method of treating a cell proliferative disorder by binding an antibody to the sample specimen are for example, classified in at least Class 424, subclass 85.8.Note that in Group IV, claim 14 is generic toaplurality of recited species which are: species (A) a neoplasm or (B) endometriosis (claim 15); species (C) a skeletal disorder (claim 16) determined by in vivo measurement (claim 17) using labeled (claim 18) compounds which are (claim 19) radioisotopes (Class 435, subclass 504); species (D) a skeletal disorder (claim 16) determined by in vivo measurement (claim 17) using labeled (claim 18) compounds which are (claim 19) luminescent compounds (Class 435, subclass 8); species (E) where the measurement is in vitro (claim 20) and the antibody is labeled (claim 21) with radioisotopes (Class 436, subclass 504); species (F) where the measurement is in vitro (claim 20) and the antibody is labeled (claim 21) with luminescent compounds such as in Class 435, subclass 8. Species A of Group IV will be examined with Group IV should applicant pay the additional fee for searching Group IV where species B through F constitute five additional species which will be searched upon payment of the requisite additional fees for each species.

In Group V, claim 23 is generic to a plurality of recited species of reagent and cell proliferative disorder which consist of species (A, claim 24) anti-GDF-5 antibody (Class 424, subclass85.8); species (B, claim 25) GDF-5 antisense polynucleotide (Class 536, subclass 23.1); species (C, claim 28)where the reagent is avector which is a colloidal dispersion (claim 29) is a liposome (claim 30) which is targeted (claims 31-37); species (D, claim 28) where the reagent

is a vector which isa (claims 38-41) virus (Class 435, subclass 235.1); species(E)wherein the disorder is a neoplasm or (F)endometriosis (claim 26); species (G) wherein the disorder is a skeletal disorder (claim 27). Species A of Group V will be examined with Group IV should applicant pay the additional fee for searching Group IV where species B through G constitute six additional species which will be searched upon payment of the requisite additional fees for each species.